

138758

STIC-Biotech/ChemLib

From: Chan, Christina
Sent: Friday, November 26, 2004 3:51 PM
To: Schnizer, Richard; STIC-Biotech/ChemLib
Subject: RE: 10/798,532 RUSH

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Schnizer, Richard
Sent: Friday, November 26, 2004 7:14 AM
To: Chan, Christina
Subject: 10/798,532 RUSH

Please authorize the following RUSH search request:

Please search the commercial and interference databases for SEQ ID NOS: 1-4 from 10/798,532. Please also search for nucleic acids that could encode the polypeptides of SEQ ID NOS: 2 and 4.

Thank you-

Richard Schnizer, Ph.D.
Patent Examiner
Art Unit 1635
Remsen 2D30
571-272-0762
Mail Box 2C18

STAFF USE ONLY

Searcher: Noble
Searcher Phone: 2-
Date Searcher Picked up: Nov 30 2004
Date Completed: 3
Searcher Prep/Rev. Time: 3
Online Time: 3

Type of Search
NA Sequence: # 4
AA Sequence: # 2
Structure: #
Bibliographic:
Litigation:
Patent Family:
Other:

Vendors and cost where applicable

STN:
DIALOG:
QUESTEL/ORBIT:
LEXIS/NEXIS:
SEQUENCE SYSTEM: COMPU
WWW/Internet:
Other(Specify):

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:30 ; Search time 5310 Seconds

(without alignments)
8646.716 Million cell updates/sec

Title: US-10-798-532-1

Perfect score: 1260

Sequence: 1 atcgaggagatccgcgtcgsc.....tgtatcaacctgtcttaa 1260

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hlc:*
4: gb_est3:*
5: gb_est3:*
6: gb_est5:*
7: gb_est5:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|-------------|
| 1 | 111.5 | 88.5 | 1763 | 3 | AK013347 | Mus muscult |
| 2 | 88.6 | 70.4 | 2846 | 3 | AK030152 | Mus muscult |
| 3 | 86.1 | 68.4 | 1079 | 4 | BM468107 | AGENCOURT |
| 4 | 723.8 | 57.4 | 917 | 4 | BM451184 | AGENCOURT |
| 5 | 624.4 | 49.6 | 1015 | 6 | BY712344 | BY712344 |
| 6 | 619.2 | 49.1 | 657 | 4 | BG547675 | 602575520 |
| 7 | 617.8 | 49.0 | 899 | 2 | BE895119 | 601436041 |
| 8 | 616.4 | 48.9 | 772 | 4 | BI562420 | 603256316 |
| 9 | 610.4 | 48.4 | 932 | 5 | BU541306 | AGENCOURT |
| 10 | 595 | 47.2 | 875 | 4 | BI259568 | 602968330 |
| 11 | 593.2 | 47.1 | 834 | 7 | CF593363 | AGENCOURT |
| 12 | 585.2 | 46.4 | 682 | 4 | BI133581 | 602997065 |
| 13 | 583.8 | 46.3 | 906 | 4 | BG386905 | 602454704 |
| 14 | 579.6 | 46.0 | 888 | 5 | BQ225071 | AGENCOURT |
| 15 | 560 | 44.4 | 602 | 7 | CO737403 | SLT02C17 |
| 16 | 558 | 44.3 | 876 | 5 | BU184346 | AGENCOURT |
| 17 | 557.6 | 44.3 | 779 | 5 | BU447831 | 603767571 |
| 18 | 550.8 | 43.7 | 844 | 4 | BG474616 | 602517423 |
| 19 | 519.8 | 41.3 | 824 | 2 | BE568528 | 601342425 |
| 20 | 516.8 | 41.0 | 823 | 7 | CO771206 | testis ES |
| 21 | 516 | 41.0 | 781 | 7 | CN232453 | WB074D11 |
| 22 | 499.4 | 39.6 | 670 | 7 | CN359673 | 170004245 |
| 23 | 498.6 | 39.6 | 717 | 5 | BU385761 | 603582650 |
| 24 | 496.8 | 39.4 | 523 | 2 | BF929410 | IL2-NT020 |

| | | | | | |
|----|-------|------|------|---|--------------------|
| 25 | 495.2 | 39.3 | 896 | 7 | CO773091 testis ES |
| 26 | 494.6 | 39.3 | 928 | 5 | BO424438 AGENCOURT |
| 27 | 494.2 | 39.2 | 1003 | 5 | BO922020 AGENCOURT |
| 28 | 482.4 | 38.3 | 716 | 7 | CF745774 UI-M-GV0- |
| 29 | 471.8 | 37.4 | 755 | 5 | BU371858 603568233 |
| 30 | 469.8 | 37.3 | 794 | 5 | BU205357 603104569 |
| 31 | 468.8 | 37.2 | 785 | 6 | CB990192 AGENCOURT |
| 32 | 468.2 | 37.2 | 824 | 4 | BI253163 602973284 |
| 33 | 465 | 36.9 | 699 | 5 | BU707284 UI-M-FR0- |
| 34 | 463.6 | 36.8 | 870 | 5 | BQ216813 AGENCOURT |
| 35 | 463.4 | 36.8 | 884 | 5 | BX281211 BX281211 |
| 36 | 460 | 36.5 | 696 | 5 | BU439278 604147528 |
| 37 | 456.6 | 36.2 | 483 | 2 | BF929411 IL2-NT020 |
| 38 | 454.2 | 36.0 | 631 | 5 | BU112717 603131421 |
| 39 | 451.8 | 35.9 | 743 | 4 | BI562440 603256324 |
| 40 | 450.6 | 35.8 | 933 | 2 | BE886040 601507192 |
| 41 | 450.4 | 35.7 | 986 | 5 | BU751496 CH4#002_E |
| 42 | 433.2 | 34.4 | 820 | 7 | CK597557 AGENCOURT |
| 43 | 427.4 | 33.9 | 672 | 5 | BX261106 BX261106 |
| 44 | 420 | 33.3 | 604 | 5 | BX261107 BX261107 |
| 45 | 413.2 | 32.8 | 570 | 5 | BU293329 604166812 |

ALIGNMENTS

| | | | | | |
|------------|--|-------------|------|--------|-----------------|
| RESULT 1 | AK013347 | 1763 bp | mRNA | linear | HTC 03-APR-2004 |
| LOCUS | AK013347 | | | | |
| DEFINITION | Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:281045404 product:kinase interacting with leukemia-associated gene (statmin), full insert sequence. | | | | |
| ACCESSION | AK013347 | GI:12850651 | | | |
| VERSION | AK013347.1 | | | | |
| KEYWORDS | HTC; CAP trapper. | | | | |
| SOURCE | Mus musculus (house mouse) | | | | |
| ORGANISM | Mus musculus | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus. | | | | |
| AUTHORS | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | | | | |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | | | | |
| JOURNAL | Genome Res. 10 (10), 1617-1630 (2000) | | | | |
| MEDLINE | 20499374 | | | | |
| PUBMED | 99279253 | | | | |
| AUTHORS | 10349636 | | | | |
| REFERENCE | 2 | | | | |
| AUTHORS | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | | | | |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | | | | |
| JOURNAL | Genome Res. 10 (10), 1617-1630 (2000) | | | | |
| MEDLINE | 20499374 | | | | |
| PUBMED | 11042159 | | | | |
| AUTHORS | 3 | | | | |
| TITLE | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. | | | | |
| TITLE | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer | | | | |
| JOURNAL | Genome Res. 10 (11), 1757-1771 (2000) | | | | |
| MEDLINE | 20530913 | | | | |
| PUBMED | 11076861 | | | | |
| AUTHORS | 4 | | | | |
| TITLE | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. | | | | |
| JOURNAL | Functional annotation of a full-length mouse cDNA collection | | | | |
| REFERENCE | Nature 409, 685-690 (2001) | | | | |
| AUTHORS | 5 | | | | |
| TITLE | The FANTOM Consortium and the RIKEN Genome Exploration Research | | | | |

Group Phase I & II Team
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1763)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,

| TITLE | JOURNAL | REFERENCE | AUTHORS |
|--|---------|----------------------------|---------|
| Group Phase I & II Team. | | | |
| Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs | | Nature 420, 563-573 (2002) | |
| 6 (bases 1 to 1763) | | | |
| Adachi, J., Atzawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kaikawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasai, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Shibata, K., Tagami, M., Tagawa, A., Takehashi, F., Tanaka, T., Teijima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Yamatsu, M. and Hayashizaki, Y. | | | |
| Direct Submission | | | |
| Submitted (10-JUL-2000) | | | |
| Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-reseegsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) | | | |

COMMENT

Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN-Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGACAGCATCCAGACGCTCTTTTCTTTTTTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGACAGCTTCGAGTTAATTAATTATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end; XhoI; 3' end; SstI. Host: SOIR.

```

FEATURES
source
location/Qualifiers
1..1763
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:2810454004"
/db_xref="Ensembl:10090"
/clone="2810454004"
/issue_type="whole body"
/clone_id="Riken full-length enriched mouse cDNA library"
/dev_stage="10, 11 days embryo"
159..1418

```

```

/ncore="unnamed protein product: kinase interacting with
leukemia-associated gene (batatmin) (MGI:1341368)
putative"
/codon_start=1
/protein_id="BAB28802.1"
/db_xref="GI:12850652"
/translation="MAGSGCWAQAEPPRFLEAFGLWQVOSRLGSSGSASYRVRCGG
TPSPGALKOPLPGTTGAASAAYGFRKERALAOLOSHRIYVLYGFTTFSP
NPPRCRLLELDVSEVELLYSHSGCKMMNIOHCARDLEAFHSHGVYADLK
PRNLTMSAENECPLIDRGLSRKQNDVYLTDDYRPAEELQNLDAQGSDFE
CTSAVDMSLGLITLLEMPSGMKLQTRYSQEMKNSAIIIDHIFSAAYVNAALPAYH
LRDILKSLHDDPGRRIIAEWALCSPPFSITFAAIDLDLWLPVYLRNLNVDVDDY
ENDEBYEDVDEKCEQKGPVSVLSLPENPGRGOVFEYANAGSKAAQKLLTGR
MFOEKFWAIFYPLSAAYKRGILYDTLL"

```

| ORIGIN | | | | |
|----------------------------|-------|----------------|-----------|--------------|
| Query March | 88.5% | Score 115; | DB 3; | Length 1763; |
| Best Local Similarity | 92.8% | Pred. No. 0; | | |
| Matches 1169; Conservative | 0; | Mismatches 91; | Indels 0; | Gaps 0; |

| | | |
|----|--|------|
| OY | ATGGCGGGATTCGGCTGGCGCTGGGGGGCGGAGCGCGCGCTTTCTGGAGCGCTTCGGG | 60 |
| Db | ATGGCGGGGATTCGGCTGGCGCTGGGGGGCGGAGCGCGCGCTTCCTGGAGCGCTTCGGG | 159 |
| OY | CGGCTGTGGCAGGTACAGAGCGCGTCTGGGTAGCGGCTTCCTCGCTCGGTGTACGGGTT | 120 |
| Db | CGGCTGTGGCAGGTACAGAGCGCGCTCGGTGTACGGGCTTCCTCGCTCGGTGTACGGGTT | 219 |
| OY | CGCTGTGGCGGCAACCTCGGCTCGCCCCCGGCGCCCTCAAGCAGTTCCTGGCGGCAAG | 180 |
| Db | CGGTGTGGCGGCAACCTCGGCTCGCCCCCGGCGCCCTCAAGCAGTTCCTGGCTTCGGGA | 279 |
| OY | ACCAACGGGGCTGGCGGCTCTGGCGCGAGTATGTTCCGCAAAAGAGGCGCGCTG | 240 |
| Db | ACCAACGGGGCTGGCGGCTCGGCGCGGAGTATGTTTCGCAAAAGAGGCGCGCTG | 339 |
| OY | GAACAGTTGCAGGGTCACAGAAACATGTGTACTTTGATGGAGTGTTCACATCACTTT | 300 |
| Db | GANCAAGTTGCAGGGTTCACAGAAACATGTGTACTTTAATGAGTCTTTACATCACTTC | 399 |
| OY | TCTCCAATGAGGCATCACGCTGTCTGTGTGTTAACTCCCGAGTACAGTGTTCGGAA | 360 |
| Db | TCTCCCATGAGGCATCACGCTGTCTGTGTAACTCCCGAGTACAGTGTTCGGAA | 459 |
| OY | TTGCTCTTATATTCCAAGTACACAGAGGTTGTTCCATGTGATGATACAGACTTGCGCCGA | 420 |
| Db | TTGCTTTATATATTCCAAGTACACAGAGGCTGTCTCATGTGATGATACAGACTGTGCCGA | 519 |
| OY | GATGTTTTGAGGGCCCTGCTTTCTTCATATAGAGGGCTATGTCCATGGGCACTCAAA | 480 |
| Db | GATGTTCTGGAGGCTCTCGCTTTCTTCATCAGGAGCTACGTTCATGCAACTCAAA | 579 |
| OY | CCACGTAAACATATTTGAGAGTGCAGAGATGATATTTTAACTCAATGACTTTGACTT | 540 |
| Db | CCACGAAACATCTGTGAGAGTGCAGAGATGAGTATTTAAGCTTATGACTTTGACTC | 639 |
| OY | AGCTTCAAAGAGGCATTCAGAGATGTAAAGTATTCAGACAGCGGTAATCGGGCTCCA | 600 |
| Db | AGCTTCAAAGAGGCATTCAGAGAGTAAAGTATTCAGACAGCGGTAATCGGCTCT | 699 |
| OY | GAACAGAGATTCGCAAAATTCCTTGGGCCACGAGCTGGCTGCAGATGTATACAGATGTACC | 660 |
| Db | GAAACAGAGCTGCAGAACTGCTTGGGCCACGAGCTGGCTGCAGATGTATACAGATGTACC | 759 |
| OY | TCAGCTGTGATCTGTGAGAGCTCAAGAAATCACTTTTACTGAAATGTTCTCAGAAATGAA | 720 |
| Db | TCAGCTGTGATCTGTGAGAGCTCGAGATCACTTTACTGAAATGTTCTCAGAGATGAG | 819 |
| OY | CTGAAACATACAGTCAAGATCTCAGAGATGGAAGGCAACAGTTCTGCTATATATGATCAC | 780 |
| Db | CTGAAACATACAGTCAAGATCTCAGAGATGGAAGGCAACAGTTCTGCTATATATGATCAT | 879 |
| OY | ATATTTGCCAGTAAAGCAGTGTGAATGCCGCAATTCAGGCTATACCTCAGAGACTT | 840 |
| Db | ATATTTGCCAGTAAAGCAGTGTGAATGCCGCAATTCAGGCTATACCTCAGAGACTT | 939 |
| OY | ATCAAAAAGCATGCTTCAATGATATCCCAAGCAGAAAGATTCCTGCTGAAATGCGATGTGCG | 900 |
| Db | ATCAAAAAGCATGCTTCAATGATGACCCGAGCAGAAAGATTCCTGCTGAGATGCGATGTGCG | 999 |
| OY | AGCCATTTCTTTAGCATTCCTTTTGGCCCTCATATTAAGAGTCTGTATGCTTCCACT | 960 |
| Db | AGCCATTTCTTTAGCATTCCTTTTGGCCCTCATATTAAGAGTCTGTATGCTTCCACT | 1059 |
| OY | CCAAGTCTAAGACTGCTGAATGTGCTGATGTATATCTTGGGAATGAAGAGAAATAT | 1020 |
| Db | CCAAGTCTCAGGCTCTTCAATGTGCTGATGTATTAACCTTGAAATGAAGAGAAATAT | 1119 |
| OY | GAACATGTTGTAGAGATGTAAAGAGAGAGTGTCAAAAAATATGAGCAGAGTGTCTCTTA | 1080 |
| Db | GAAATGTTGTAGAGATGTGAAAGAGAGAGTGTCAAGAAATATGAGCAGAGTGTCTCTG | 1179 |
| OY | CTTGTTCCAAAGAAAATCCTGGACAGAGCAAGACTTTTGTGAGTATGCAAAATGCTGCT | 1140 |

| | | |
|---|----------|---|
| | TITLE | Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., |
| | JOURNAL | Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., |
| | | Muramatsu, W. and Hayashizaki, Y. |
| | COMMENT | Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedic Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/. Location/Qualifiers |
| | FEATURES | location |
| | Source | 1 . 2846 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:4933401D07" /db_xref="taxon:10090" /c clone="4933401D07" /sex="male" /tissue type="testis" /clone_id="RIKEN_full-length enriched mouse cDNA library" /dev_stage="adult" 238 . 1230 /note="unnamed protein product; kinase interacting with leukemia-associated gene (statmin) (MGI:MGI:1341908) putative" /codon_start=1 /protein_id="BAC26812.1" /db_xref="GI:26326137" /translation="MTLYGVFTTFHFPNPSRCLLELDVSVSELLYSHGQCSNM MIHQCARVDLEALAFLHHBGYVHADLPKRNILASNEBCPKLIDFGISFKREGNDVXY IORDGRAPAEALONCIAOGSDPTECAVLWSIGILLMEFSMKKHVRSOEB WRNSSAIIIDHTFASKAVYMAATPAHYLRPLISMHDHPGRIPAMALCSPEFIIP FAHIIDLVMLPPVPIRLNLVDDDTLENDEVEDVDVEEDVKESCCQKPVSLLPYKE NPERGVFVRVIANAGSKAAQKLTGTSMRGKFVVFPTPLSAFYKGYLYQLT" 2833 . 2828 /note="putative" 2846 /note="putative" |
| | CDS | polya_signal polya_site ORIGIN Query Match 70.4%; Score 886.6; DB 3; Length 2846; Best Local Similarity 93.1%; Pred. No. 1.8e-248; Matches 928; Conservative 0; Mismatches 69; Indels 0; Gaps 0; |
| D | | 264 CATTGGACCTTTGTAATGGAAGTGTTTACAATCCACTTTTCCTCCAATGTGCCATACCGCTG 323 |
| D | | 234 CTTCATCACTTTATATGAGAGCTTTTACCATAACCTTCTCCCAATGTGCATCACCGCTG 293 |
| O | | 324 TCTGTTGCTTGAACCTCGTAGATGCACTGTTTGGAAATGCTCTTATATTCAGATCA 383 |
| D | | 294 TCTGCTGCTGAACTCTCGAATGCACTGTTTGGAAATGCTTTATATTCAGATCA 353 |
| O | | 384 GGGTGTTCATGATGATGATATACAGCATTCGCGCCGAGATGTTTTGGAGGCCCTTGTCTT 443 |
| D | | 354 GGGGTGCTCATATGATGATATACAGCACTGTGCCGAGATGTTCTGGAGGCCCTCGCTTT 413 |
| O | | 444 TCTTCATCATGAGGGCTATATGTCATGCGGACCTCAACCAACGTAACATATTTGGAATGC 503 |
| D | | 414 TCTTCACCAATGAGGGCTACGTCATGACGACCTCAACCAACGAAACATCTCTGTGGAATGC 473 |
| O | | 504 AGGAATGATGTTTTAACTCACTTGAAGCTTTGAGCTTGAAGGCAATATGAGA 563 |
| D | | 474 GGGAATGAGTGTTTAAGCTTATATGACTTTGAGCTGAGCTTCAAGAGGCAATATGAGA 533 |

QY 1047 GGAGTGCAGAAATATGACACAGTGGATCTTACTTGTTCAGAAAGAAATCTGACAG 1106
| | | | |
DB 781 GGAGTGCAGAAATATGACACAGTGGATCTTACTTGTTCAGAAAGAAATCTGACAG 840
| | | | |
QY 1107 AGGACAGATC-TTGTGTGAGTATGCAAAATGCTG-GTGAATCCAAAGCTGGCAGAAATT 1163
| | | | |
DB 841 AGGACAGATCTTTTGTGTGAGTATGCAAAATGCTGNGTGAATCCAAAGCTGGCAGAAATT 899
| | | | |
RESULT 4
BM451184 917 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
DEFINITION 5', mRNA sequence.
ACCESSION BM451184
VERSION BM451184.1 GI:18500224
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 917)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1M12121 row: j column: 24
High quality sequence stop: 718.
Location/Qualifiers
1. 917
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5495159"
/issue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 57.4%; Score 723.8; DB 4; Length 917;
Best Local Similarity 99.7%; Pred. No. 8.7e-201;
Matches 725; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 534 TGGACTTACCTTAAAGAGGCAATCAGATGTAAGTATATTCACAGACAGGGGATCG 553
| | | | |
DB 21 TGGGCTTAGCTTAAAGAGGCAATCAGATGTAAGTATATTCACAGACAGGGGATCG 80
| | | | |
QY 594 GGCTCCAGAGCAGAAATGCAAAATGCTGGCCAGGCTGGCCTGCAGAGTATCAGA 653
| | | | |
DB 81 GGCTCCAGAGCAGAAATGCAAAATGCTGGCCAGGCTGGCCTGCAGAGTATCAGA 140
| | | | |
QY 654 ATGTACCTCAGCTGTGTATCTGAGACCTAGGAATCATTTTACGGAATGTTCTCAG 713
| | | | |
DB 141 ATGTACCTCAGCTGTGTATCTGAGACCTAGGAATCATTTTACGGAATGTTCTCAG 200
| | | | |
QY 714 AATGAAATGAAACATTCAGTCTGAGATTCAGAAATGGAAGCAACAGTTCTCTATTA 773
| | | | |
DB 201 AATGAAATGAAACATTCAGTCTGAGATTCAGAAATGGAAGCAACAGTTCTCTATTA 260
| | | | |
QY 774 TGATCACAATTTTGCAGTAAGCAGTGTGATGCGCAATTCAGCCTATCACTTAAG 833
| | | | |

DB 261 TGATCACAATTTTGCAGTAAGCAGTGTGATGCGCAATTCAGCCTATCACTTAAG 320
| | | | |
QY 834 AAGCTTATCAAAAGCAGTCTTCATGATGATTCAGACAGAAATTCCTGTAATGCG 893
| | | | |
DB 321 AAGCTTATCAAAAGCAGTCTTCATGATGATTCAGACAGAAATTCCTGTAATGCG 380
| | | | |
QY 894 ATTGTGACAGCCATTCCTTAGCATTCCTTTGCCCTCATATTTAGATCTGTCATGCT 953
| | | | |
DB 381 ATTGTGACAGCCATTCCTTAGCATTCCTTTGCCCTCATATTTAGATCTGTCATGCT 440
| | | | |
QY 954 TCCCACTCCAGTCTTAAGACTGCTGATGCTGATGATGATTTATCTTTGGAAATGAGA 1013
| | | | |
DB 441 TCCCACTCCAGTCTTAAGACTGCTGATGCTGATGATGATTTATCTTTGAAATGAGA 500
| | | | |
QY 1014 GGAATATGAAAGATGTTGTGAAAGATGTAATAAGAGAGTCTCAAAATATGACCCAGTGT 1073
| | | | |
DB 501 GGAATATGAAAGATGTTGTGAAAGATGTAATAAGAGAGTCTCAAAATATGACCCAGTGT 560
| | | | |
QY 1074 ATCTCTACTTGTTCAGAAAGAAATCTCTGACAGAGACCAAGCTTTGTGATGATGCAAA 1133
| | | | |
DB 561 ATCTCTACTTGTTCAGAAAGAAATCTCTGACAGAGACCAAGCTTTGTGATGATGCAAA 620
| | | | |
QY 1134 TCTGTGATGATTCAGAAAGCTGCGCAGAAATTAAGCTGATGATGATGATGGAAGTT 1193
| | | | |
DB 621 TCTGTGATGATTCAGAAAGCTGCGCAGAAATTAAGCTGATGATGATGGAAGTT 680
| | | | |
QY 1194 TGTGTGATGATTCAGAAAGCTGCGCAGAAATTAAGCTGATGATGATGGAAGTT 1253
| | | | |
DB 681 TGTGTGATGATTCAGAAAGCTGCGCAGAAATTAAGCTGATGATGGAAGTT 740
| | | | |
QY 1254 GCTTTAA 1260
| | | | |
DB 741 GCTTTAA 747
| | | | |
RESULT 5
BY712344 1015 bp mRNA linear EST 17-DEC-2002
LOCUS BY712344
DEFINITION BY712344 RIKEN full-length enriched, 10, 11 days embryo whole body
MUS musculus cDNA clone 281045404 5', mRNA sequence.
ACCESSION BY712344
VERSION BY712344.1 GI:27123621
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1015)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oshio, N., Saito, R., Suzuki, H., Yamana, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Kiyosawa, C., Gotohori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brucic, V.,
Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Fraser, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, B.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, M., King, B.L., Konagaya, A.,
Kurachkin, I.V., Lee, Y., Lemhard, B., Lyons, P.A., Maglott, D.R.,
Malta, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavani, W.J., Petrea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shinada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wyshew-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itch,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL 22354683
MEDLINE 12466851
PUBMED 12466851

CONTACT: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, url:http://genome.gsc.riken.jp/

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
Fukuda,S., Hahizume,W., Hayashida,K., Hirozane,T., Hori,F.,
Imocani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
Kondo,S., Kono,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
Takeda,Y., Waki,K., Watabiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submision
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES

source
Location/Qualifiers
1..1015
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2810454004"
/tissue_type="whole body"
/dev_stage="10, 11 days embryo"
/clone_lib="RIKEN full-length enriched, 10, 11 days embryo
whole body"

ORIGIN

Query Match 49.6%; Score 624.4; DB 6; Length 1015;
Best Local Similarity 90.2%; Pred. No. 1.4e-171;
Matches 679; Conservative 0; Mismatches 72; Indels 2; Gaps 1;
1 ATGGCGGGAATCCGGCTGGCGCTGGGCGCGAGACCCCGCTTTCTTGAGAGGCTTCGGG 60
159 ATGGCGGGAATCCGGCTGGCGCTGGGCGCGAGACCCCGCTTTCTTGAGAGGCTTCGGG 218
61 CGGCTGTGCAAGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGCTGTGAGGTT 120
219 CGGCTGTGCAAGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGCTGTGAGGTT 278
121 CGCTGTGCGCAACCTGTGCTGCGCCCGCGGCGCTTCAAGAGATTCTTGCGCGAGGA 180
279 CGGTGTGCGGTAACCGGCGCTGCGCCCGCGGCGCTTCAAGAGATTCTTGCGCGAGGA 338
181 ACCACCGGGGCTGCGGCTCTGCGCGGAGTATGCTTTCGCAAGAAGAGGGGCGGCTG 240
|||||

Db 339 ACCACCGGGGCTGCGGCTTCGGCCCGGAGATATGCTTCCGCAAGAAGAGGGCGGCTG 398
Qy 241 GAACAGTTGCAAGGTACAGAAACATCGACTTTGATAGAGTGTATTAATCACTTT 300
Db 399 GANCAATTGCAAGGTACAGAAACATCGACTTTGATAGAGTGTATTAATCACTTT 458
Qy 301 TCTCCAAATGTCATACAGCTGTCTGTGCTTGAACCTCTGATGTCAGTGTTCGGAA 360
Db 459 TCTCCAAATGTCATACAGCTGTCTGTGCTTGAACCTCTGATGTCAGTGTTCGGAA 518
Qy 361 TTGCTTTATATTCAGTACAGAGGTTGTCAGTGTGATGATACAGATTTGCCCGGA 420
Db 519 TTGCTTTATATTCAGTATAGAGGCTGCTCAATGTGATATACAGCATGTGCCGA 578
Qy 421 GATGTTTGAAGGCTTCTTCTTCATCATAGAGGCTATGTCATCGAGCTCAAA 480
Db 579 GATGTTTGAAGGCTTCTTCTTCATCATAGAGGCTATGTCATCGAGCTCAAA 638
Qy 481 CCACGTATATTTGTGAGTGCAGAGAAATGAATGTTTAACTCAATGACTTTGACTT 540
Db 639 CCACGTATATTTGTGAGTGCAGAGAAATGAATGTTTAACTCAATGACTTTGACTT 698
Qy 541 AACTTCAAGAAGGCAATCAGATGTAAAGTATTTACAGACAGAGGATGCGGCTCA 600
Db 699 AACTTCAAGAAGGCAATCAGATGTAAAGTATTTACAGACAGAGGATGCGGCTCT 758
Qy 601 GAAGCAGATTGCAAAATGCTTGGCCGAGCTGCTGCAGAGTATCAGAAATGTAAC 660
Db 759 GAAGCAGATTGCAAAATGCTTGGCCGAGCTGCTGCAGAGTATCAGAAATGTAAC 818
Qy 661 TAGCTGTGATCTGTGAGCTTACGAAATCATTTACTGGAATGTTCTCAGAAATGAA 720
Db 819 TAGCTGTGATCTGTGAGCTTACGAAATCATTTACTGGAATGTTCTCAGAAATGAA 878
Qy 721 CTGAACATACAGTCAAGATCTCAGAAATGGAAG 753
Db 879 CTGAACATACAGTCAAGATCTCAGAAATGGAAG 909
|||||

RESULT 6
BG547675 657 bp mRNA linear EST 04-APR-2001
LOCUS 602575520P1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703645 5,
DEFINITION
mRNA sequence.
ACCESSION BG547675
VERSION BG547675.1 GI:13546340
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: CLOUTEC Laboratories, Inc.
cDNA Library Preparation: CLOUTEC Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10M1542 row: 9 column: 06
High quality sequence stop: 657.

FEATURES

source
Location/Qualifiers
1..657
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4703645"
/lab_host="DH10B (TI phage-resistant)"


```

Db      541 ATTCTGCTGAATGAGCATTTGACGCCCATCTTATGACATTCCTTTTGCCCTCATATT 600
QY      937 |||||
Db      601 GAGATCTGCTCATGCTTCC 621

RESULT 8
B1562420      772 bp      mRNA      linear      EST 05-SEP-2001
B1562420      60326316P1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
B1562420      mRNA sequence.
B1562420      B1562420
B1562420.1    GI:15449747
EST.
Homo sapiens (human)
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      NIH-MGC http://img.ncbi.nih.gov/
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgabbs-r@mail.nih.gov
               Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
               cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
               Toshikuni and Piero Carninci (RIKEN)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L14M1755 row: P column: 03
               High quality sequence scop: 727.
               Location/Qualifiers
               1..772
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5298674"
               /lab_host="DH10B"
               /lab_host="NIH_MGC_97"
               /note="Organ: testis; Vector: pBluescript (modified
               pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
               (BspGag); Oligo-dT primed using primer
               5'-TTTTTTTTTTTTTTN-3', size-selected for average
               insert size 2.2 kb and normalized for R01 5. This is a
               primary library enriched for full-length clones and
               constructed using the Cap-trapper method (Carninci, in
               preparation). Library constructed by M. Brownstein
               (NHGRI/NHGRI, National Institutes of Health). Note: this is
               a NIH_MGC library."

ORIGIN
Query Match      48.9%; Score 616.4; DB 4; Length 772;
Best Local Similarity 99.4%; Pred. No. 2.9e-169;
Matches 628; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY      274 TTGTATGAGAGTGTTCATCATCTTTCTCCAAATGTCGATCAGCTGTCTGTGCTT 333
Db      108 TTGTATGAGAGTGTTCATCATCTTTCTCCAAATGTCGATCAGCTGTCTGTGCTT 167

QY      334 GAACCTCTGATGCTGTTGGGAAATGCTCTTATTCAGTACACGAGGTGTGTC 393
Db      168 GAACCTCTGATGCTGTTGGGAAATGCTCTTATTCAGTACACGAGGTGTGTC 227

QY      394 ATGTGATGATAGACATTCGCGCCGAGATGTTTGGAGGCGCTGTCTTCTTCATCAT 453
Db      228 ATGTGATGATAGACATTCGCGCCGAGATGTTTGGAGGCGCTGTCTTCTTCATCAT 287

QY      454 GAGGGCTATGTCATGCGGAGCTCAACGATGATATTTGGAGTGCAAGAAATGAA 513

```

```

Db      288 GAGGGCTATGTCATGCGGAGCTCAACGATGATATTTGGAGTGCAAGAAATGAA 347
QY      514 TGTTTAACTGATGACTTTTGGACTTACCTTCAAGAGGCAATGAGATGAAATGAT 573
Db      348 TGTTTAACTGATGACTTTTGGACTTACCTTCAAGAGGCAATGAGATGAAATGAT 407

QY      574 ATTGACAGACAGCGGATGCGGCTCCGAGAGGAAATTTGCCAAATTTGCGCCAGCT 633
Db      408 ATTGACAGACAGCGGATGCGGCTCCGAGAGGAAATTTGCCAAATTTGCGCCAGCT 467

QY      634 GCGCTGCAAGAGATTCAGAAATGATCTCAGCTGTGTGATCTGTGAGCCTTAGAATCAT 693
Db      468 GCGCTGCAAGAGATTCAGAAATGATCTCAGCTGTGTGATCTGTGAGCCTTAGAATCAT 527

QY      694 TTAATGAAATGTTCTCAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 753
Db      528 TTAATGAAATGTTCTCAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 587

QY      754 GCACACAGTTCTGCTATTTATGATCA-TATTTGCGAGTAAAGCAGTGTGATGCGCC 812
Db      588 GCACACAGTTCTGCTATTTATGATCA-TATTTGCGAGTAAAGCAGTGTGATGCGCC 647

QY      813 AATTCAGGCTTACCTTAAGACCTTTATCAAAAGCATGCTTATGATTCAGACAG 872
Db      648 AATTCAGGCTTACCTTAAGACCTTTATCAAAAGCATGCTTATGATTCAGACAG 707

QY      873 AAGAAATTCCTGTAATGAGCATTTGTGAGCC 904
Db      708 AAGAAATTCCTGTAATGAGCATTTGTGAGCC 739

```

```

RESULT 9
B0541306      932 bp      mRNA      linear      EST 13-SEP-2002
B0541306      AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
LOCUS          5', mRNA sequence.
DEFINITION
ACCESSION      B0541306
VERSION        B0541306.1 GI:22851747
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      NIH-MGC http://img.ncbi.nih.gov/
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgabbs-r@mail.nih.gov
               Tissue Procurement: DCTD/DRP
               cDNA Library Preparation: Rubin Laboratory
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Agencourt Bioscience Corporation
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1CM2765 row: O column: 18
               High quality sequence scop: 631.
               Location/Qualifiers
               1..932
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:6572970"
               /issue_type="carcinoma, cell line"
               /lab_host="DH10B (phage-resistant)"
               /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
               Site 2: EcoRI; cDNA made by oligo-dT priming.
               Directionally cloned into EcoRI/XhoI sites using the
               following 5' adaptor: GGCACGAG(G). Library constructed by
               Ling Hong in the laboratory of Gerald M. Rubin (University
               of California, Berkeley) using Zap-cDNA synthesis kit

```

ORIGIN (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 48.4%; Score 610.4; DB 5; Length 932;
Best Local Similarity 93.9%; Pred. No. 1.8e-167;
Matches 690; Conservative 0; Mismatches 37; Indels 8; Gaps 5;

QY 1 ATGGCGGAGATCCGGCTGCGCTGGGGGCGGAGCCCGCGCTTTTCTGGAGGCGCTTCGGG 60
DB 149 ATGGCGGAGATCCGGCTGCGCTGGGGGCGGAGCCCGCGCTTTTCTGGAGGCGCTTCGGG 208
QY 61 CGGCTGGAGGAGTACAGAGCCGCTGGAGGAGGCGCTTCGGCTGGAGTATCCGGGTT 120
DB 209 CGGCTGGAGGAGTACAGAGCCGCTGGAGGAGGCGCTTCGGCTGGAGTATCCGGGTT 268
QY 121 CGGCTGGAGGAGTACAGAGCCGCTGGAGGAGGCGCTTCGGAGAGTTCCTGGCGCAGGA 180
DB 269 CGGCTGGAGGAGTACAGAGCCGCTGGAGGAGGCGCTTCGGAGAGTTCCTGGCGCAGGA 328
QY 181 ACCACCGGGGCTGCGCTGCGCGCGAGATATGTTTCCGCAAGAGAGGGCGGCTG 240
DB 329 ACCACCGGGGCTGCGCTGCGCGCGAGATATGTTTCCGCAAGAGAGGGCGGCTG 388
QY 241 GAACAGTGGAGGAGTACAGAAATCATGTGATTTGATGAGTGTACATCACTTT 300
DB 389 GAACAGTGGAGGAGTACAGAAATCATGTGATTTGATGAGTGTACATCACTTT 448
QY 301 TCTCAATGTGCATCAGCTGCTGTTGCTTGAATCTCGATGTCAGTGTTCGAA 360
DB 449 TCTCAATGTGCATCAGCTGCTGTTGCTTGAATCTCGATGTCAGTGTTCGAA 508
QY 361 TTGCTCTTATATTCAGTCCAGAGGTTTTCATGATGATGATACAGATTCGCCCGA 420
DB 509 TTGCTCTTATATTCAGTCCAGAGGTTTTCATGATGATGATACAGATTCGCCCGA 568
QY 421 GATGTTTGGAGGCGCTGCTTCTTCATCATGAGGCGATGTCATGCGAGCT-CAA 479
DB 569 GATGTTTGGAGGCGCTGCTTCTTCATCATGAGGCGATGTCATGCGAGCT-CAA 628
QY 480 ACCACGTAATATGAGTGCAGAGTGAATGTTTAACTCATTCATCTTGGAC- 538
DB 629 ACCACGTAATATGAGTGCAGAGTGAATGTTTAACTCATTCATCTTGGAC- 688
QY 539 TTGCTCTTATATTCAGTCCAGAGGTTTTCATGATGATGATACAGATTCGCCCGA 598
DB 689 TTGCTCTTATATTCAGTCCAGAGGTTTTCATGATGATGATACAGATTCGCCCGA 748
QY 599 CAG-AAGCAGATTTGCAAAATGCTTGGCCGAGGCG-TCGCTGAGAGTATCAGATG 656
DB 749 CAGAAACAGAAATTTGCAAAATGCTTGGCCGAGGCGTTCGCAAGTATCAGAAATG 808
QY 657 TACCT---CAGCTGTGATCTGTGAGCCTAGGAATCATTTTACCTGGAATGTTCTCAG 712
DB 809 TACCTCAGCTGTGATCTGTGAGCCTAGGAATCATTTTACCTGGAATGTTCTCAG 868
QY 713 GAATGAACCTGAAC 727
DB 869 CAGGGAATGGAAC 883

RESULT 10
LOCUS B1259568 875 bp mRNA linear EST 17-JUL-2001
DEFINITION 602968330F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5108094 5',
mRNA sequence.
ACCESSION B1259568
VERSION B1259568.1 GI:14817032
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 875)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/HLN at:
<http://image.llnl.gov>
Plate: L14M1261 row: k column: 07
High quality sequence stop: 722.
Location/Qualifiers
1. 875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5108094"
/issue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1b="NIH_MGC_12"
/note="Organ: Cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

FEATURES

source

1. 875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5108094"
/issue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1b="NIH_MGC_12"
/note="Organ: Cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 47.2%; Score 595; DB 4; Length 875;
Best Local Similarity 91.6%; Pred. No. 5.9e-163;
Matches 809; Conservative 0; Mismatches 50; Indels 24; Gaps 16;

QY 297 CTTTCTCCAAATGTGCATCAGCTGCTGTTGCTTGAATCTCGATGTCAGTGTTC 356
DB 1 CTTTCTCCAAATGTGCATCAGCTGCTGTTGCTTGAATCTCGATGTCAGTGTTC 59
QY 357 GGAATGCTCTTATATTCAGTCCAGAGGTTTTCATGATGATGATACAGATTCGCC 416
DB 60 GGAATGCTCTTATATTCAGTCCAGAGGTTTTCATGATGATGATACAGATTCGCC 417
QY 417 CCGAGATGTTTGGAGGCGCTGCTTCTTCATCATGAGGCGTATGTCATGCGAGCT 476
DB 118 CCGAGATGTTTGGAGGCGCTGCTTCTTCATCATGAGGCGTATGTCATGCGAGCT 477
QY 477 CAAACCACTTAATATGAGTGCAGAGATGAATGTTTAACTCATTCATCTTGG 536
DB 176 CAAACCACTTAATATGAGTGCAGAGATGAATGTTTAACTCATTCATCTTGG 537
QY 537 ACTTACCTTCAAGAGGCAATCAGATGTAATGTTTATTCAGACAGAGGATATGGGC 596
DB 234 ACTTACCTTCAAGAGGCAATCAGATGTAATGTTTATTCAGACAGAGGATATGGGC 597
QY 597 TCCAGAGCAGATTTCAAAATGCTTGGCCGAGGCTGCGCTGAGAGTATCAGAAATG 656
DB 294 TCCAGAGCAGATTTCAAAATGCTTGGCCGAGGCTGCGCTGAGAGTATCAGAAATG 657
QY 657 TACCTCAGCTGTGATCTGTGAGCCTAGGAATCATTTTACTGGAATGTTCTCAGAAAT 716
DB 354 TACCTCAGCTGTGATCTGTGAGCCTAGGAATCATTTTACTGGAATGTTCTCAGAAAT 717
QY 717 GAAACTGAAACATACAGTCTAGATCTTACAGAAATGGAAGGCAACAGTTCTGCTATTGGA 776
DB 414 GAAACTGAAACATACAGTCTAGATCTTACAGAAATGGAAGGCAACAGTTCTGCTATTGGA 777
QY 777 TCATATTATTTGCAGTGAAGCAGTGTGATGCGGCAATTCAGACTTACCTTAAGAGA 836
DB 474 TCATATTATTTGCAGTGAAGCAGTGTGATGCGGCAATTCAGACTTACCTTAAGAGA 837
QY 837 CCTTATCAAAAGCATCTTCATGATGATCCAGCAGAGAAATTCCTGTGAATGGCAAT 896
DB 534 CCTTATCAAAAGCATCTTCATGATGATCCAGCAGAGAAATTCCTGTGAATGGCA-T 592

| | | | |
|----|------|--|------|
| QY | 897 | GTGAGCCCAATCTTTAGCATTCCTTTGGCCCTCAATTAAAGATC-AGTCATAGCTTC | 955 |
| Db | 593 | GTGAGCCCAATCTTTAGCATTCCTTTGGCCCTCATATTGAAGTCGGGTCAAGCTTC | 652 |
| QY | 956 | CCAATCCAGTC-CTAAGACTGCTGAAATGTGCTGGATGATGATTAATCTTGGGAATGAA-G | 1012 |
| Db | 653 | CCAATCCAGGGGCTAAAGACTGCTGAAATGTGCTGGATGATGATTAATCTTGGGAATGAAAG | 712 |
| QY | 1013 | AGGAATATGAAG-ATGTGTAGAAAG--ATGTAAAGAGAGTGTCAAAATATATGAGACCAG | 1065 |
| Db | 713 | AGGAATATGAGAATGTTGTGAGAGAGATGTAAAGAGAGATGTCAACCAATTTGGACCA | 772 |
| QY | 1070 | TGTATCTCTAATTGTT-----CCAAAGAAAATCTGSCAGAGGACA-AGTCTTTGTT | 1122 |
| Db | 773 | GGTGTATTCCTCTACTTGTGTCCAAAAGGAAAAATCTCGGGCGAGACACAGCTTTGTGA | 832 |
| QY | 1123 | G-AGTATGCAAAAT-GCTGTGATTCCAAAGCGGCGAGAAATTT | 1163 |
| Db | 833 | GAAATATGCAAAATGTTGGGAGATTCCAAAGATGCGCGAAATTT | 875 |

| RESULT 11 | LOCUS | DEFINITION |
|--------------------|-------------|-------------------------|
| CF593363 | 834 bp | linear |
| CF593363 | 834 bp | EST 26-SEP-2003 |
| AGENCOURT 15622279 | NIH_MGC_147 | Homo sapiens cDNA clone |
| IMAGE:30531268 | 5', | mRNA sequence. |

| | | |
|----------|----------------------|-------------|
| VERSION | CF593363.1 | GI:36346826 |
| KEYWORDS | EST. | |
| SOURCE | Homo sapiens (human) | |
| ORGANISM | Homo sapiens | |

| REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|--------------------|--|---|---------|---------|
| 1 (bases 1 to 834) | NIH-MGC | http://mgc.nci.nih.gov/ . | | |
| | National Institutes of Health, Mammalian Gene Collection (MGC) | | | |
| | Unpublished (1999) | | | |
| | Contact: Daniela S. Gerhard, Ph.D. | | | |

Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLNL at:
<http://image.lnl.gov>
Plate: NDAm621 row: m column: 05
High quality sequence stop: 660.

```

FEATURES
source
Location/Qualifiers
1..834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30531268"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_idb="NIH_MGC_147"
/notes="Organ: placenta; Vector: pBluescriptR; Site_1:
all: XhoI; Site_2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to R07 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carinci, in
preparation). Library constructed by M. Brownstein
(NIH/NGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

```

Query Match 47.1%; Score 593.2; DB 7; Length 834;
Best Local Similarity 97.6%; Pred. No. 2e-162;
Matches 645; Conservative 0; Mismatches 9; Indels 7; Gaps 4;

| | | | |
|----|-----|--|-----|
| QY | 1 | ATGGCGGGATTCGGGCTGCGCTTGGGGCGCGAGCCGCGCGTCTTTCTGAGAGGCTTCGGG | 60 |
| Db | 155 | ATGGCGGGATTCGGGCTGCGCTTGGGGCGCGAGCCGCGCGTCTTTCTGAGAGGCTTCGGG | 214 |
| QY | 61 | CGGCTGTGGCAGGTATCAGAGCCGCTGTGGGTAGCGGCTCTCCGGCTGGGTATTCGGGTT | 120 |
| Db | 215 | CGGCTGTGGCAGGTATCAGAGCCGCTGTGGGTAGCGGCTCTCCGGCTGGGTATTCGGGTT | 274 |
| QY | 121 | CGCTGTGCGGCAACCTTGCTCGCCGCCCGGCGCCTCAAGACATTCCTTGCCCGCAGAA | 180 |
| Db | 275 | CGCTGTGCGGCAACCTTGCTCGCCGCCCGGCGCCTCAAGACATTCCTTGCCCGCAGAA | 334 |
| QY | 181 | ACCAACCGGGGCTGCGGCTCTGCGCGGAGTAGTGTTCGCAAAAGAGGGCGGCGCTG | 240 |
| Db | 335 | ACCAACCGGGGCTGCGGCTCTGCGCGGAGTAGTGTTCGCAAAAGAGGGCGGCGCTG | 394 |
| QY | 241 | GAACAGTTGCAGGGTCAACGAAACATGTGTACTTTGATGGAAGTATTACAAATCCACTTT | 300 |
| Db | 395 | GAACAGTTGCAGGGTCAACGAAACATGTGTACTTTGATGGAAGTATTACAAATCCACTTT | 454 |
| QY | 301 | TCTCCAATGATGCAATCAGCGCTGTCTGTGCTTGAACTCCCGAATGCAGTGTTTGGAA | 360 |
| Db | 455 | TCTCCAATGATGCAATCAGCGCTGTCTGTGCTTGAACTCCCGAATGCAGTGTTTGGAA | 514 |
| QY | 361 | TTGCTCTTAATATTCAGATCACAGGGTGTTCATGTAATGATATACAGCATTTGGCCGA | 420 |
| Db | 515 | TTGCTCTTAATATTCAGATCACAGGGTGTTCATGTAATGATATACAGCATTTGGCCGA | 574 |
| QY | 421 | GATGTTTTGAGAGCCCTTGCTTTCTTCATCATAGAGGCTATGTTCATTCGGAGACTCAA | 480 |
| Db | 575 | GATGTTTTGAGAGCCCTTGCTTTCTTCATCATAGAGGCTATGTTCATTCGGAGACTCAA | 634 |
| QY | 481 | CCAGCTAACATATTTGTGAGTGCAGAGATGAATGTTTTAACTCATTTGACTTTGGAATT | 540 |
| Db | 635 | CCAGCTAACATATTTGTGAGTGCAGAGATGAATGTTTTAACTCATTTGACTTTGGAATT | 694 |
| QY | 541 | A-GCTTCAAGAGAGGCATCAGATGTAAATATATTCAACAGACGGGTATC-GGGCTC | 598 |
| Db | 695 | AGGCTTCAAGAGAGGCATCAGATGTAAATATATTCAACAGACGGGTATTCGGGGCTC | 754 |
| QY | 599 | CAGAAAGCAATTCGAAATTTGCTT---GGGCCAGGCT--GGCTGCAGAGTATATACGA | 653 |
| Db | 755 | CAGAAAGCAATTCGAAATTTGCTTGGGCCCGCAGGCTTGCCCTGCAGAGTATATACGA | 814 |
| QY | 654 | A 654 | |
| Db | 815 | A 815 | |

| | |
|------------|--|
| RESULT | 12 |
| LOCUS | B1333581 |
| DEFINITION | B1333581 682 bp mRNA linear EST 30-JUL-2001 |
| ACCESSION | 60297065F1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:5139184 5', |
| VERSION | mRNA sequence. |
| KEYWORDS | B1333581 B1333581 GI:15018238 |
| SOURCE | EST. |
| ORGANISM | Homo sapiens (human) |
| | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |
| REFERENCE | 1 (bases 1 to 682) |
| AUTHORS | NIH-MGC http://mgc.nci.nih.gov/ . |
| TITLE | National Institutes of Health, Mammalian Gene Collection (MGC) |
| JOURNAL | Unpublished (1999) |
| COMMENT | Contact: Robert Strausberg, Ph.D. Email: gga@nci.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. |

CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL1342 row: j column: 17
High quality sequence stop: 674.

FEATURES

source

1..682
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5139184"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

ORIGIN

Query Match 46.4%; Score 585.2; DB 4; Length 682;
Best Local Similarity 97.3%; Pred. No. 4.1e-160;
Matches 648; Conservative 0; Mismatches 13; Indels 5; Gaps 5;

```

QY 1 ATGGCGGGGATCCGGCTGGCGGGGCGGAGACCGCGCTTTTGTGAGAGGCTTCGGG 60
DB 19 ATGGCGGGGATCCGGCTGGCGGGGCGGAGACCGCGCTTTTGTGAGAGGCTTCGGG 78
QY 61 CGGCTGTGCGAGGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGCTGATCCGGGTT 120
DB 79 CGGCTGTGCGAGGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGCTGATCCGGGTT 138
QY 121 CGGCTGTGCGAGACCTTGCTGCGCCCGCGCGCTTCAAGAGTTCTTCCGCGAGAA 180
DB 139 CGGCTGTGCGAGACCTTGCTGCGCCCGCGCGCTTCAAGAGTTCTTCCGCGAGAA 198
QY 181 ACCACGGGGGCTGGCGGCTCTGCGCGCGAGTAGTTTCCGAAAGAGGGGCGGCTG 240
DB 199 ACCACGGGGGCTGGCGGCTCTGCGCGCGAGTAGTTTCCGAAAGAGGGGCGGCTG 258
QY 241 GAACAGTTCAGAGGTACAGAAACATCGTGAATTTGATGAGTGTTCATCACTTT 300
DB 259 GAACAGTTCAGAGGTACAGAAACATCGTGAATTTGATGAGTGTTCATCACTTT 318
QY 301 TCTCCAAATGTGCATCAGCTGTGTGCTTGAACCTCTGATGTCAAGTGTTCGAA 360
DB 319 TCTCCAAATGTGCATCAGCTGTGTGCTTGAACCTCTGATGTCAAGTGTTCGAA 378
QY 361 TTGCTCTTATATTCAGTACCGAGGTGTTCATGTGATGATTAAGCATTTGCCGCA 420
DB 379 TTGCTCTTATATTCAGTACCGAGGTGTTCATGTGATGATTAAGCATTTGCCGCA 438
QY 421 GATGTTTGGAGGCTTGTCTTTTTCATCATGAGGCTATGCTCATGCGAGCTCAAA 480
DB 439 GATG-TTTGGAGGCTTGTCTTTTTCATCATGAGGCTATGCTCATGCGAGCTCAAA 497
QY 481 CCAAGTAAATTTGAGAGTGAAGATGATTAATTTAACTCATTTGAGACTT 540
DB 498 CCAAGTAAATTTGAGAGTGAAGATGATTAATTTAACTCATTTGAGACTT 557
QY 541 AGCTTCAAGAGGAGCAATCA-GGATGTAAGTATATTCAGAGG-AAGGATATCCGGGCTC 598
DB 558 AGCTTCAAGAGGAGCAATCAAGATGTAAGTATATTCAGAGG-AAGGATATCCGGGCTC 617
QY 599 CAGAGCAGATTTGCAAAATTTGCTGGCCAGAGCTGGCTGAGAGTGAATTA-CAGATGT 657
DB 618 CAGAGCAGATTTGCAAAATTTGCTGGCCAGAGCTGGCTGAGAGTGAATTA-CAGATGT 676
QY 658 ACCTCA 663
DB 677 ACCTCA 682

```

RESULT 13

BG386905

LOCUS

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

602454704P1 NIH_MGC_15 Homo sapiens CDNA clone IMAGE:4583065 5',
mRNA sequence.
BG386905
BG386905.1 GI:13280454
EST.
Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 906)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM1306 row: 9 column: 02
High quality sequence stop: 668.
Location/Qualifiers
1..906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4583065"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_15"
/note="Organ: colon; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"

FEATURES

source

ORIGIN

Query Match 46.3%; Score 583.8; DB 4; Length 906;
Best Local Similarity 96.1%; Pred. No. 1.2e-159;
Matches 664; Conservative 0; Mismatches 17; Indels 10; Gaps 6;

```

QY 570 GTATATTCAGACAGACGGGATATCGGGCTTCAGAGCAGAAATTCGCTTGGCCCA 629
DB 2 GTATATTCAGACAGACGGGATATCGGGCTTCAGAGCAGAAATTCGCTTGGCCCA 61
QY 630 GGCTGGCTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 689
DB 62 GGCTGGCTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121
QY 690 CATTATTCAGAAATGTTCTCAGAGATGAACTGAAACATACAGTCAGATCTCAGAAATG 749
DB 122 CATTATTCAGAAATGTTCTCAGAGATGAACTGAAACATACAGTCAGATCTCAGAAATG 181
QY 750 GAAGGCAACAGTTGCTATTATTTATGATCATATTTTGGCAATAAGAGTGTGAATGC 809
DB 182 GAAGGCAACAGTTGCTATTATTTATGATCATATTTTGGCAATAAGAGTGTGAATGC 241
QY 810 CGCAATTCAGAGCTTACCTTAAGAGACCTTTCAAAAGCAGTCTTCAGATGATCCAAAG 869
DB 242 CGCAATTCAGAGCTTACCTTAAGAGACCTTTCAAAAGCAGTCTTCAGATGATCCAAAG 301
QY 870 CAGAGAAATTCCTGCTGAATGAGCATTTGTCAGCCCATTTCTTACATTCCTTTGGCC 929

```

```

Db      302 CAGAAAGATTCCTGCTGAATGCGATGTGACAGCCCATCTTTAGCATCTTTGGCCCC 361
Qy      930 TCATTTGAAAGATCTGGTCAATGCTTCCCACTCCAGTGTAAAGCTGCTGAATGTGCTGA 989
Db      362 TCATTTGAAAGATCTGGTCAATGCTTCCCACTCCAGTGTAAAGCTGCTGAATGTGCTGA 421
Qy      990 TCATGATTTATCTTGGGATGAGAGAGAAATGAAAGATGTTGTAAGATGTAAGAGGA 1049
Db      422 TGATGATTTATCTTGGGATGAGAGAGAAATGAAAGATGTTGTAAGATGTAAGAGGA 481
Qy      1050 GTGTCAAAAATATGACCAAGTGTATCTTACTTTGTCAGAGAAAATCTGCGAGAGG 1109
Db      482 GTGTCAAAAATATGACCAAGTGTATCTTACTTTGTCAGAGAAAATCTGCGAGAGG 539
Qy      1110 ACAAGCTTTGTTGATGATGCAAAATGCTGATGATCCCAAGCTGCGAGAAAATTAAGTAC 1169
Db      540 ACAAGCTTTGTTGATGATGCAAAATGCTGATGATCCCAAGCTGCGAGAAAATTAAGTAC 598
Qy      1170 TGGAGAGATGTTGATGAGGAAGTTGTTGCTGCTACATTCACCCGCTGAGTCCATCA 1229
Db      599 TGGAGG--AGTTGATGAGGAAGTTGTTGCTGCTACATTCACCCGCTGAGTCCATCA 651
Qy      1230 GAGGGATATCTGTATCAAACTTCTCTTTAA 1260
Db      652 AGAGGATATCTGTATCAAACTGCGCTTTAA 682

```

```

RESULT 14
LOCUS      BQ225071      888 bp      mRNA      linear      EST 02-MAY-2002
DEFINITION AGENCOURT 732333 NIH_MGC_70 Homo sapiens cDNA IMAGE:6016796
5', mRNA sequence.
ACCESSION  BQ225071
VERSION    BQ225071.1  GI:20406471
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strauberg, Ph.D.
            Email: csabbs-remail.nih.gov
            Tissue Procurement: ARCC
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: LLMK13215 row: a column: 21
            High quality sequence stop: 581.
            Location/Qualifiers
                1..888
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:6016796"
                /issue_type="epithelioid carcinoma"
                /lab_host="DH10B (phage-resistant)"
                /clone_1lb="NIH MGC 70"
                /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.1 Kb. Library constructed by Life
                Technologies."

```

```

FEATURES
    source
        1..888
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:6016796"
        /issue_type="epithelioid carcinoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_1lb="NIH MGC 70"
        /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
        Average insert size 1.1 Kb. Library constructed by Life
        Technologies."

```

```

ORIGIN
Query Match      46.0%; Score 579.6; DB 5; Length 888;
Best Local Similarity 97.9%; Pred. No. 2e-158;
Matches 619; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

```

```

Qy      1  ATGGCGGATCCGCTGCGCTGCGGCGGAGCCGCGCTTTCTGAGGCTTCGGG 60
Db      149  ATGGCGGATCCGCTGCGCTGCGGCGGAGCCGCGCTTTCTGAGGCTTCGGG 208
Qy      61  CGGCTGTGGCAGGTACAGAGCCGCTCTGGGTAGCGCTCTCGGCTCGGTATTCGGGTT 120
Db      209  CGGCTGTGGCAGGTACAGAGCCGCTCTGGGTAGCGCTCTCGGCTCGGTATTCGGGTT 268
Qy      121  CGGCTGTGGGCAACCTGCGCTGCGGCGGCGCTCAAGCAGTTCTTGCGCGCAGGA 180
Db      269  CGGCTGTGGGCAACCTGCGCTGCGGCGGCGCTCAAGCAGTTCTTGCGCGCAGGA 328
Qy      181  ACCACCGGGGCTGCGGCTCTGCGGCGGAGTATGTTTCGCAAGAGAGGCGGCGCTG 240
Db      329  ACCACCGGGGCTGCGGCTCTGCGGCGGAGTATGTTTCGCAAGAGAGGCGGCGCTG 388
Qy      241  GAACAGTTGCGAGGTACAGAAACATGAGACTTTGTATGAGAGTGTTCATTCACCTT 300
Db      389  GAACAGTTGCGAGGTACAGAAACATGAGACTTTGTATGAGAGTGTTCATTCACCTT 448
Qy      301  TCTCCAAATGTGCGATCAGCTGTCTGTGCTTGAACCTCGGATGTCAGTGTTCGGA 360
Db      449  TCTCCAAATGTGCGATCAGCTGTCTGTGCTTGAACCTCGGATGTCAGTGTTCGGA 508
Qy      421  GATGTTTGGAGG--CCCTTGCTTTTCTTATCATAGAGGCTATGTCACGCCGACTCA 479
Db      569  GATGTTTGGAGGCGCCCTTGCTTTTCTTATCATAGAGGCTATGTCACGCCGACTCA 628
Qy      480  ACCACGTAACATATTTGAGAGTGCAGAGATGTAATTTTAACTCATGACTTTGAGACT 539
Db      629  ACCACGTAACATATTTGAGAGTGCAGAGATGTAATTTTAACTCATGACTTTGAGACT 688
Qy      597  TCCAGAGCAGAAATTCGCAAAATTTGCTTGCGCC 628
Db      749  TCCAGAGCAGAAATTCGCAAAATTTGCTTGCGCC 780

```

```

RESULT 15
LOCUS      CO737403      602 bp      mRNA      linear      EST 29-JUL-2004
DEFINITION SILT02c17116f1 squirrel testis library 1 Spermophilus lateralis
5', mRNA sequence.
ACCESSION  CO737403
VERSION    CO737403.1  GI:50824673
KEYWORDS   EST.
SOURCE      Spermophilus lateralis (golden-mantled ground squirrel)
ORGANISM   Spermophilus lateralis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Scurinae;
            Spermophilus.

```

```

REFERENCE  1 (bases 1 to 602)
AUTHORS   Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
            Rogers,J. and Cossins,A.R.
            Microarray analysis of transcriptional changes during hibernation
            in the golden mantled ground squirrel, Spermophilus lateralis
            Unpublished (2004)
            Contact: Andrew R. Cossins
            Laboratory for Environmental Gene Regulation
            University of Liverpool
            School of Biological Sciences, The Biosciences Building, Crown
            Street, Liverpool, United Kingdom, L69 7ZB
            Tel: +44(0)151-795-4510
            Fax: +44(0)151-795-4431
            Email: cossins@lly.ac.uk

```

Vector has been trimmed from this EST.

Plate: 17 row: i column: 16

Seq primer: pf1c T7 (5'-AATACGACTCACTATAGG-3')

High quality sequence stop: 602.

Location/Qualifiers

1..602

/organism="Spermophilus lateralis"

/mol_type="mRNA"

/db_xref="taxon:76772"

/clone="17116"

/sex="Male & female"

/tissue type="Testis"

/dev stage="Adult"

/lab_host="R. coli Electromax DH10B"

/clone_lib="Squirrel testis library 1"

/note="Vector: pFLC; Site_1: SalI GTCGAG; Site_2: BamHI

GGATCC; Normalized and subtracted cDNA library prepared

from testis of hibernating and summer animals"

ORIGIN

Query Match 44.4%; Score 560; DB 7; Length 602;

Best Local Similarity 95.8%; Pred. No. 1e-152;

Matches 575; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
QY 624 GGGCCAGGCTGGCTGCAGAGTATACAGAAATGATACCTGAGCTGTGATCTGGAGCCT 683
DB 3 GGGCCAGGCTGGCTGCAGAGTATACAGAAATGATACCTGAGCTGTGATCTGGAGCCT 62
QY 684 AGGAATCATTTTACTGGAAATGTTCTCAGAAATGAACTGAAACATACAGTCAGATCTCA 743
DB 63 AGGAATTTATTTTACTGGAAATGTTCTCAGAAATGAACTGAAACATACAGTTAGATCTCA 122
QY 744 GGAATGGAAGGCAACAGTTCTCTATTTATTTGATTCACATTTTGGCCATTAACAGTGGT 803
DB 123 GGAATGGAAGGCAAAATGTTCTCTATTTATTTGATTCACATTTTGGCCATTAACAGTGGT 182
QY 804 GAATGCCGCAATTCAGCCTATCACCTTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 863
DB 183 GAATGCCGCAATTCAGCCTATCACCTTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 242
QY 864 TCCAGCAGAAAGATTCCTGCTGAAATGGCAATGTGAGCCCATTCCTTAGCATTCCTTT 923
DB 243 TCCAGCAGAAAGATTCCTGCTGAAATGGCAATGTGAGCCCATTCCTTAGCATTCCTTT 302
QY 924 TGCCCTCATTTGAAGATCTGTCATGCTTCCCATCTCCAGTCTAAGACTGCTGAATGT 983
DB 303 TGCCCTCATTTGAAGATTTGGTGAATGCTTCTACTCCAGTCTAAGATTTGCTCAATGT 362
QY 984 GCTGATGATGATTTATCTTGGAAATGAAGAAATATGAATGATTTGTAAGATGTAA 1043
DB 363 CTTGATGATGATTTATCTTGGAAATGAAGAAATATGAATGATTTGTAAGATGTAA 422
QY 1044 AGAGGAGTGCAGAAATATGACCAAGTGTATCTTACTTTCAGAAAGAAATCTCTGG 1103
DB 423 AGAGGAGTGCAGAAATATGACCAAGTGTATCTTACTTTCAGAAAGAAATCTCTGG 482
QY 1104 CAGAGGCAAGTCTTTGTTGATGATGCAAAATGCTGGTATTCAGAAAGCTGCGCAAAAT 1163
DB 483 CAGAGGCAAGTCTTTGTTGATGATGCAAAATGCTGGTATTCAGAAAGCTGCGCAAAAT 542
QY 1164 ACTGACTGAAAGATGTTTGAAGGAAATTTGTTGGCTACATTCACCCGCTGAGTGC 1223
DB 543 ACTGACTGAAAGATGTTTGAAGGAAATTCGTTGGCTACATTCACCCGCTGAGTGC 602
```

Search completed: November 29, 2004, 17:46:23
Job time : 5316 secs

This Page Blank (uspto)


```

Db      181 ACCACCGGGGCTCGGCTCTGCGCCGAGATGTTTCGCAAGAGAGGGGGGCTG 240
Qy      241 GAACAGTTGAGGGTCAACAGAAACATCGTACCTTTGATGAGATGTTTACATCCACTT 300
Db      241 GAACAGTTGAGGGTCAACAGAAACATCGTACCTTTGATGAGATGTTTACATCCACTT 300
Qy      301 TCTCAAAATGTCATCAGCTGTCTGTTGCTTGAATCCTGATGTCAGTGTTCGGA 360
Db      301 TCTCAAAATGTCATCAGCTGTCTGTTGCTTGAATCCTGATGTCAGTGTTCGGA 360
Qy      361 TTGCTCTTATATTCAGTCAACAGGGTGTTCATGTGATGATACAGATTCGCCCA 420
Db      361 TTGCTCTTATATTCAGTCAACAGGGTGTTCATGTGATGATACAGATTCGCCCA 420
Qy      421 GATGTTTGGAGGCCCTTGCTTTTCTTCATCAGAGGGCTATGTCCATGCGACCTCAA 480
Db      421 GATGTTTGGAGGCCCTTGCTTTTCTTCATCAGAGGGCTATGTCCATGCGACCTCAA 480
Qy      481 CCACGTAAACATATTTGAGAGTGCAGAGAAATGATGTTTAACTCATTTGACTT 540
Db      481 CCACGTAAACATATTTGAGAGTGCAGAGAAATGATGTTTAACTCATTTGACTT 540
Qy      541 AGCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGCGGATATCGGGCTCA 600
Db      541 AGCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGCGGATATCGGGCTCA 600
Qy      601 GAAGCAGAAATTGCAAAATTTGCTGGCCAGGCTGGCTGAGATGATACAGAAATGAC 660
Db      601 GAAGCAGAAATTGCAAAATTTGCTGGCCAGGCTGGCTGAGATGATACAGAAATGAC 660
Qy      661 TCAGCTGTGATCTGTGAGCTAGAGATCATTTTACTGGAATGTTTCTCAGAAATGAA 720
Db      661 TCAGCTGTGATCTGTGAGCTAGAGATCATTTTACTGGAATGTTTCTCAGAAATGAA 720
Qy      721 CTGAAACATACATCAGATCTCAGAAATGAGCAACAGTCTCTATATTTATGATCAC 780
Db      721 CTGAAACATACATCAGATCTCAGAAATGAGCAACAGTCTCTATATTTATGATCAC 780
Qy      781 ATATTGGCCAGTAAAGAGTGAATGCGCAATTCAGCTATACCTTAAGAGACTT 840
Db      781 ATATTGGCCAGTAAAGAGTGAATGCGCAATTCAGCTATACCTTAAGAGACTT 840
Qy      841 ATCAAAAGCATGCTTCATGATGATCCAAAGCAGAAATTCCTGCTGAATGCGATTTG 900
Db      841 ATCAAAAGCATGCTTCATGATGATCCAAAGCAGAAATTCCTGCTGAATGCGATTTG 900
Qy      901 AGCCCATTTCTTATGATCTCTTTTGGCCCTCATATTTGAAGATCTGGTCACTGCCACT 960
Db      901 AGCCCATTTCTTATGATCTCTTTTGGCCCTCATATTTGAAGATCTGGTCACTGCCACT 960
Qy      961 CCAGTCTAAGATGCTGGAATGCTGATGATGATTTATCTGGGATGAAGAGAAAT 1020
Db      961 CCAGTCTAAGATGCTGGAATGCTGATGATGATTTATCTGGGATGAAGAGAAAT 1020
Qy      1021 GAAGATGTTGTAGAAAGATGTAAGAGAGAGTGTCAAAATATATGACCAAGTATCTTA 1080
Db      1021 GAAGATGTTGTAGAAAGATGTAAGAGAGAGTGTCAAAATATATGACCAAGTATCTTA 1080
Qy      1081 CTGTGTTCCAAAGAAATCTGTGCAAGAGCAAGTCTTTGTGATGATGAAATGCTGGT 1140
Db      1081 CTGTGTTCCAAAGAAATCTGTGCAAGAGCAAGTCTTTGTGATGATGAAATGCTGGT 1140
Qy      1141 GATTCCAAAGCTCGCAGAAATTAATGATGATGAGAGATGTTTGAATGGAAGTTTGTG 1200
Db      1141 GATTCCAAAGCTCGCAGAAATTAATGATGATGAGAGATGTTTGAATGGAAGTTTGTG 1200
Qy      1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGCTTTAA 1260
Db      1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGCTTTAA 1260

```

RESULT 2

```

AX262521
LOCUS      AX262521      1260 bp      DNA      linear      PAT 26-OCT-2001
DEFINITION Sequence 12 from Patent WO0173050.
ACCESSION AX262521
VERSION    AX262521.1  GI:16511408
KEYWORDS
SOURCE
ORGANISM   Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1. Meyer, R. 3714, 16742, 23546, and 13887 novel protein kinase molecules and
  uses thereof. Patent: WO 0173050-A 12 04-OCT-2001;
  Millennium Pharmaceuticals, Inc. (US)
FEATURES
  source
    location/Qualifiers
      1..1260
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
ORIGIN
Query Match      99.7%; Score 1256.8; DB 6; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy      1 ATGGCGGGGATCCGGGCTGCGGCTGGGGCGGAGCCGCGCTTTCTGGAGGCTTCGGG 60
Db      1 ATGGCGGGGATCCGGGCTGCGGCTGGGGCGGAGCCGCGCTTTCTGGAGGCTTCGGG 60
Qy      61 CGGCTGTGACAGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTCGGTGTATCGGGTT 120
Db      61 CGGCTGTGACAGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTCGGTGTATCGGGTT 120
Qy      121 CGCTGTGCGGCAACCTTGCTGCGCCCGCGCGCTTCACAGATCTTGGCCGACAGA 180
Db      121 CGCTGTGCGGCAACCTTGCTGCGCCCGCGCGCTTCACAGATCTTGGCCGACAGA 180
Qy      181 ACCACCGGGGCTGCGGCTCTGTCGCGCGAGTATGTTTCCGAAAGAGAGGCGGCTG 240
Db      181 ACCACCGGGGCTGCGGCTCTGTCGCGCGAGTATGTTTCCGAAAGAGAGGCGGCTG 240
Qy      241 GAACAGTTGACAGGCTCACAGAAACATCGTACTTTGTATGAGTGTTCATTCACACTT 300
Db      241 GAACAGTTGACAGGCTCACAGAAACATCGTACTTTGTATGAGTGTTCATTCACACTT 300
Qy      301 TCTCCAAATGTCATCAGCTGTCTGTTGCTTGAATCCTGATGTCAGTGTTCGGA 360
Db      301 TCTCCAAATGTCATCAGCTGTCTGTTGCTTGAATCCTGATGTCAGTGTTCGGA 360
Qy      361 TTGCTCTTATATTCAGTCAACAGGGTGTTCATGTGATGATACAGATTCGCCCA 420
Db      361 TTGCTCTTATATTCAGTCAACAGGGTGTTCATGTGATGATACAGATTCGCCCA 420
Qy      421 GATGTTTGGAGGCCCTTGCTTTTCTTCATCAGAGGGCTATGTCCATGCGACCTCAA 480
Db      421 GATGTTTGGAGGCCCTTGCTTTTCTTCATCAGAGGGCTATGTCCATGCGACCTCAA 480
Qy      481 CCACGTAAACATATTTGAGAGTGCAGAGAAATGATGTTTAACTCATTTGACTT 540
Db      481 CCACGTAAACATATTTGAGAGTGCAGAGAAATGATGTTTAACTCATTTGACTT 540
Qy      541 AGCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGCGGATATCGGGCTCA 600
Db      541 AGCTTCAAGAGGCAATCAGATGTAAGTATATTCAGACAGCGGATATCGGGCTCA 600
Qy      601 GAAGCAGAAATTGCAAAATTTGCTGGCCAGGCTGGCTGAGATGATACAGAAATGAC 660
Db      601 GAAGCAGAAATTGCAAAATTTGCTGGCCAGGCTGGCTGAGATGATACAGAAATGAC 660
Qy      661 TCAGCTGTGATCTGTGAGCTAGAGATCATTTTATGGAATGTTTCTCAGAAATGAA 720

```

```

Db      TCAGCTGTTAGCTGTGAGAGCTTGAAGATCATTTTACTGGAATGTTCTCAGGAATGAA 720
Qy      721 CTGAAACATACGTCAGATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATGATCAC 780
Db      721 CTGAAACATACGTCAGATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATGATCAC 780
Qy      781 ATATTGCGCAGAAAGAGTGTGAAATGCGCAATTCAGCCTTATCACTTAAGACCTT 840
Db      781 ATATTGCGCAGAAAGAGTGTGAAATGCGCAATTCAGCCTTATCACTTAAGACCTT 840
Qy      841 ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCCTGCTGAAATGCGCATTTGTC 900
Db      841 ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCCTGCTGAAATGCGCATTTGTC 900
Qy      901 AGCCCATTTCTTAAAGATTCCTTTTGGCCCTCATATTGAAAGATCTGTCATGCTCCACT 960
Db      901 AGCCCATTTCTTAAAGATTCCTTTTGGCCCTCATATTGAAAGATCTGTCATGCTCCACT 960
Qy      961 CCAGTGTCTAAGATGCTGTAATGCTGATGATGATTAATCTTGGGAATGAGAGGAATAT 1020
Db      961 CCAGTGTCTAAGATGCTGTAATGCTGATGATGATTAATCTTGGGAATGAGAGGAATAT 1020
Qy      1021 GAAGATGTTGTGAAGATGTAAAGAGAGTGTCAAAAATATGAGCAAGTGTATCTCTTA 1080
Db      1021 GAAGATGTTGTGAAGATGTAAAGAGAGTGTCAAAAATATGAGCAAGTGTATCTCTTA 1080
Qy      1081 CTTGTTCCAAAGGAAATCTCTGCGCAGAGACAAGTCTTTTGTGATGATGCAAAATGCTGCT 1140
Db      1081 CTTGTTCCAAAGGAAATCTCTGCGCAGAGACAAGTCTTTTGTGATGATGCAAAATGCTGCT 1140
Qy      1141 GATTCCAAAGCTGCGCAGAGAAATTAAGTCTGAGAGGATGTTGATGGAAGTTTGTGTG 1200
Db      1141 GATTCCAAAGCTGCGCAGAGAAATTAAGTCTGAGAGGATGTTGATGGAAGTTTGTGTG 1200
Qy      1201 GCTACATTTCTACCCGCTGAGTGTCTTCAAGAGGAGATCTGTATCAAACTTGTCTTAA 1260
Db      1201 GCTACATTTCTACCCGCTGAGTGTCTTCAAGAGGAGATCTGTATCAAACTTGTCTTAA 1260

```

```

RESULT 3
AX262519      2598 bp      DNA      linear      PAT 26-OCT-2001
LOCUS      AX262519
DEFINITION      Sequence 10 from Patent WO0173050.
ACCESSION      AX262519
VERSION      AX262519.1 GI:16511406
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Meyers, R.
TITLE      3714, 16742, 23546, and 13887 novel protein kinase molecules and
JOURNAL      Patent: WO 0173050-A 10 04-OCT-2001;
FEATURES
SOURCE      Millennium Pharmaceuticals, Inc. (US)
location/Qualifiers
1..2598
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
268..1527
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10238.1"
/db_xref="GI:16511407"
/translation="MAGSGCAGWAGAPPRFLAFLRLMQVQSHLSSGSSASVYRVACG
NGSPGALKQFLPGTTGAASAAYGRKRALBOLQGRNVTLYGVTHFSP
NPSRCLLELDIVSRLSYSHOGSMWIOHCAQVYLAFLHREGVHADIK
PNHILMSAENECFKLIDRFLSFKGNQVXYQTQDGYRAPZELQNCALQALQSDTE
CTSAVDLWSLGIILLEMFSGLKHTVRSQEWKANSALIDHIFASKAVVNAISAYH
LRDLIKSMHDDPSRRIPALMALCSPFSLPPAPHIEDLVMLPTVLRLNLVLDLYL
ENBEEDVEDVKEKCKYGVVSLVLPKNRGRGVVEYANAGDSKAAQKLTGR

```

```

ORIGIN
MFDGKFVATFVPLSAVRYLYQLL"
Query Match 99.7%; Score 1256.8; DB 6; Length 2598;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy      1 ATGGGGGATCCGGTGTGCGCTGGGGGCGGAGCGCGCGCTTTCTGGAGCCCTTCCGG 60
Db      268 ATGGGGGATCCGGTGTGCGCTGGGGGCGGAGCGCGCGCTTTCTGGAGCCCTTCCGG 327
Qy      61 CGGCTGTGGCAGGTACAGAGCGCTGTGGGTAGCGGCTTCTCCGCTCGGTATCGGGTT 120
Db      328 CGGCTGTGGCAGGTACAGAGCGCTGTGGGTAGCGGCTTCTCCGCTCGGTATCGGGTT 387
Qy      121 CGCTGTGGGGAACCTTGCGTCCGCCCGCGCGCTTCAAGCAGTTTTCGCCGACAGA 180
Db      388 CGCTGTGGGGAACCTTGCGTCCGCCCGCGCGCTTCAAGCAGTTTTCGCCGACAGA 447
Qy      181 ACCACCGGGGCTGGGGGCTGTGGCGCGAGTATGGTTCCGCAAAAGAGGGGCGCGCTG 240
Db      448 ACCACCGGGGCTGGGGGCTGTGGCGCGAGTATGGTTCCGCAAAAGAGGGGCGCGCTG 507
Qy      241 GAACAGTTGCAAGGTCAAGAAACATCGTGAATTGTATGAGAGTGTTCACATCCACTT 300
Db      508 GAACAGTTGCAAGGTCAAGAAACATCGTGAATTGTATGAGAGTGTTCACATCCACTT 567
Qy      301 TCTCCAAATGTCGCAATCAGCGCTGTCTGTGCTTGAATCTCTGATGTGATGTTTGGAA 360
Db      568 TCTCCAAATGTCGCAATCAGCGCTGTCTGTGCTTGAATCTCTGATGTGATGTTTGGAA 627
Qy      361 TTGCTCTTATATTCAGATCACAGGGTGTTCATGATGATGATACAGATTTGCGGCCGA 420
Db      628 TTGCTCTTATATTCAGATCACAGGGTGTTCATGATGATGATACAGATTTGCGGCCGA 687
Qy      421 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
Db      688 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 747
Qy      481 CCACGTAAATATTTGTGAGATGTCAGAGATGAATGTTTAAATCTGATGATCTTGGACTT 540
Db      748 CCACGTAAATATTTGTGAGATGTCAGAGATGAATGTTTAAATCTGATGATCTTGGACTT 807
Qy      541 AGCTTCAAGAGGAAATCAGAGATGAATGAATATTCAGACAGAGGATTCGGGCTCCA 600
Db      808 AGCTTCAAGAGGAAATCAGAGATGAATGAATATTCAGACAGAGGATTCGGGCTCCA 867
Qy      601 GAAGCAGATTTGCAAAATTTGCTGGCCCGAGCTGCGCTGAGAGTATACAGATGTACC 660
Db      868 GAAGCAGATTTGCAAAATTTGCTGGCCCGAGCTGCGCTGAGAGTATACAGATGTACC 927
Qy      661 TCAGCTGTTAGCTGTGAGAGCTTGAAGATCATTTTACTGGAATGTTCTCAGGAATGAA 720
Db      928 TCAGCTGTTAGCTGTGAGAGCTTGAAGATCATTTTACTGGAATGTTCTCAGGAATGAA 987
Qy      721 CTGAAACATACGTCAGATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATGATCAC 780
Db      988 CTGAAACATACGTCAGATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATGATCAC 1047
Qy      781 ATATTGCGCAGTAAAGAGTGTGAAATGCGCAATTCAGCCTTATCACTTAAGACCTT 840
Db      1048 ATATTGCGCAGTAAAGAGTGTGAAATGCGCAATTCAGCCTTATCACTTAAGACCTT 1107
Qy      841 ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCCTGCTGAAATGCGCATTTGTC 900
Db      1108 ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCCTGCTGAAATGCGCATTTGTC 1167
Qy      901 AGCCCATTTCTTAAAGATTCCTTTTGGCCCTCATATTGAAAGATCTGTCATGCTCCACT 960
Db      1168 AGCCCATTTCTTAAAGATTCCTTTTGGCCCTCATATTGAAAGATCTGTCATGCTCCACT 1227
Qy      961 CCAGTGTCTAAGATGCTGTAATGCTGATGATGATTAATCTTGGGAATGAGAGGAATAT 1020

```

Db 1228 CCAGTGTCTAAGCTGCTGATGTCTGATGATGATATCTTGAGATGAGAGAAATAT 1287
Qy 1021 GAAGATGTTGTAGAGATGTAAAGAGAGAGTGTCAAAAATATGACCAAGTGTATCTTA 1080
Db 1288 GAGATGTTGTAGAGATGTAAAGAGAGAGTGTCAAAAATATGACCAAGTGTATCTTA 1347
Qy 1081 CTTGTTCCAAAGGAAATCTGTGACAGAGCAAGTCTTTGTTAGTATGCAATGCTGGT 1140
Db 1348 CTTGTTCCAAAGGAAATCTGTGACAGAGCAAGTCTTTGTTAGTATGCAATGCTGGT 1407
Qy 1141 GATTCCAAAGCTGCGGAGAAATTAATGATCTGGAAGATGTTTATGGAAGTTGTTG 1200
Db 1408 GATTCCAAAGCTGCGGAGAAATTAATGATCTGGAAGATGTTTATGGAAGTTGTTG 1467
Qy 1201 GCTACATTCACCCGCTGAGTGTCTTACAGAGGGGATATCTGTATCAAACTTCTTAA 1260
Db 1468 GCTACATTCACCCGCTGAGTGTCTTACAGAGGGGATATCTGTATCAAACTTCTTAA 1527

RESULT 4
HSA536197
LOCUS HSA536197 2909 bp mRNA linear PRI 11-JAN-2003
DEFINITION Homo sapiens mRNA for KIS protein.
ACCESSION AJ536197
VERSION AJ536197.1 GI:27657360
KEYWORDS KIS protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1
TITLE Bieche, I., Manceau, V., Curmi, P. A., Laurendon, I., Lachkar, S., Leroy, K., Vidaud, D., Sobel, A. and Maucuer, A.
JOURNAL Quantitative RT-PCR reveals a ubiquitous but preferentially neural expression of the KIS gene in rat and human
REFERENCE 2 (bases 1 to 2909)
AUTHORS Maucuer, A.
TITLE Direct Submision
JOURNAL Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du fer a moulin, 75005 Paris, FRANCE

FEATURES
source
Location/Qualifiers
1..2909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:IMAG998M213680"
/db_xref="taxon:9606"
/clone="I.M.A.G.E. Consortium CloneID 1455548"
/issue_type="2 pooled tumors (clear cell type)"
1..2909
/gene="KIS"
146..1405
/gene="KIS"
/codon_start=1
/product="KIS protein"
/protein_id="CAD60192.1"
/db_xref="GI:27657361"
/translation="MAGSGCAGMGABPREFLEAFGRMLQVNOYRLSGSSASVYRVRCG
NGSPGAKKQPLPGTTGAASAAEYGRKEBALEOLQGHNIYTLKGVFLHSP
NPSRCLLELDVSYSELLYSHOGCSMMKIQHCARVLEALATLHHEGVYHDLK
PRTILWSAENECFKLIDFGLSFEKGNQDVKTOTDGRAPBAELQCLQAGLQSDTE
CTSAVDLWSIGIILLEMFGMKLKHTRVQEWKANSALIDHIFASKAVVNAIPAYH
LBDLILKSMLEHDPSSRIPAEMALCSPPFSIPAPAHIEDLVMLPTVLRILNVLDYLL
ENEEYEDVDEKCECKYKQPVSLVLEPNRGQVFVEYVANNAGDSKAAQCLTLGR
MDDGRVAVTFYPLSAVKRGYLQTLT"

ORIGIN

Query Match 99.7%; Score 1256.8; DB 9; Length 2909;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCGGATCCGGCTGCGCTGGGGCGCGAGCCCGCGGCTTTTCTGAGAGCCTTCGGG 60
|||

Db 146 ATGGCGGATCCGGCTGCGCTGGGGCGCGAGCCCGCGGCTTTTCTGAGAGCCTTCGGG 205
Qy 61 CGAGTGTGACAGATACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTATCGGGT 120
Db 206 CGAGTGTGACAGATACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTATCGGGT 265
Qy 121 CGCTGTGCGGACCTTGGCTTCGCCCCCGGCGCTTCAGACAGTTCTTTCGCGCAGGA 180
Db 266 CGCTGTGCGGACCTTGGCTTCGCCCCCGGCGCTTCAGACAGTTCTTTCGCGCAGGA 325
Qy 181 ACCACCGGGGCGCGGCTTCGCCCCCGGAGTATGTTTCCGAAAGAGGGCGGCGCTG 240
Db 326 ACCACCGGGGCGCGGCTTCGCCCCCGGAGTATGTTTCCGAAAGAGGGCGGCGCTG 385
Qy 241 GAACAGTTGACAGGATCACAGAAACATGATGATTTGATGAGTGTATTAATCCACTTT 300
Db 386 GAACAGTTGACAGGATCACAGAAACATGATGATTTGATGAGTGTATTAATCCACTTT 445
Qy 301 TCTCCAAATGTGCAATCACGCTGTCTGTTGCTTGAACCTCTGATGTCACTGTTCCGAA 360
Db 446 TCTCCAAATGTGCAATCACGCTGTCTGTTGCTTGAACCTCTGATGTCACTGTTCCGAA 505
Qy 361 TTGCTCTTATATTCAGATCACAGGGTGTTCATGATGATGATTAACAGATTCGCGCGGA 420
Db 506 TTGCTCTTATATTCAGATCACAGGGTGTTCATGATGATGATTAACAGATTCGCGCGGA 565
Qy 421 GATGTTTGAAGGCGCTTCTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
Db 566 GATGTTTGAAGGCGCTTCTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 625
Qy 481 CCAAGTACATATTTGTGAGTGCAGAGAAAGAAATGTTTAACTCATGACTTGGACTT 540
Db 626 CCAAGTACATATTTGTGAGTGCAGAGAAAGAAATGTTTAACTCATGACTTGGACTT 685
Qy 541 ACCTTCAAAGAGGCAATGAGATGTAAGTATTTCAAGACAGAGGGATCGGGCTCA 600
Db 686 ACCTTCAAAGAGGCAATGAGATGTAAGTATTTCAAGACAGAGGGATCGGGCTCA 745
Qy 601 GAAGCAGATTCGAAATTCCTTGGCCGAGGCTGCGCTGACAGATGATACAGATGTACC 660
Db 746 GAAGCAGATTCGAAATTCCTTGGCCGAGGCTGCGCTGACAGATGATACAGATGTACC 805
Qy 661 TCAAGCTTGTATCTGTGAGGCTTAGAATCATTTTATCTGAAATGTTCTCAGGAATGAAA 720
Db 806 TCAAGCTTGTATCTGTGAGGCTTAGAATCATTTTATCTGAAATGTTCTCAGGAATGAAA 865
Qy 721 CTGAAACATACAGTCAAGATCTCAGGATGAGAGGCAACAGTTCTGCTATTTATGATCAC 780
Db 866 CTGAAACATACAGTCAAGATCTCAGGATGAGAGGCAACAGTTCTGCTATTTATGATCAC 925
Qy 781 ATATTTCCAGTAAAGCAGTGTGATGATGCGCAATTCAGCCTTATCAGCTAAGAGACTT 840
Db 926 ATATTTCCAGTAAAGCAGTGTGATGATGCGCAATTCAGCCTTATCAGCTAAGAGACTT 985
Qy 841 ATCAAAAGATGCTTCATGATGATTCAGACAGAAAGAAATTCCTGCTGAAATGCAATTTGTC 900
Db 986 ATCAAAAGATGCTTCATGATGATTCAGACAGAAAGAAATTCCTGCTGAAATGCAATTTGTC 1045
Qy 901 AGCCCATTTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATCTGATCTGCTCCACT 960
Db 1046 AGCCCATTTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATCTGATCTGCTCCACT 1105
Qy 961 CCAGTGTCTAAGCTGCTGATGATGCTGATGATGATATCTTGGAGATGAAGAGAAATAT 1020
Db 1106 CCAGTGTCTAAGCTGCTGATGATGCTGATGATGATATCTTGGAGATGAAGAGAAATAT 1165
Qy 1021 GAAGATGTTGTGAAGATGTAAAGAGAGATGTCAAAAATATGACCAAGTGTATCTCTTA 1080
Db 1166 GAAGATGTTGTGAAGATGTAAAGAGAGATGTCAAAAATATGACCAAGTGTATCTCTTA 1225
Qy 1081 CTTGTTCCAAAGGAAATCTGTGACAGAGCAAGTCTTTGTTGAGATGCAATGCTGGT 1140
Db 1226 CTTGTTCCAAAGGAAATCTGTGACAGAGCAAGTCTTTGTTGAGATGCAATGCTGGT 1285

| | | | |
|----------------------------|--|---|-----------------|
| QY | 1141 | GATTCCAAAGCTGGCGACAAGTAATTTCTGACTGAGAGAGAGTTTGAAGGAGACTTTGTTGTC | 1200 |
| Db | 1286 | GATTCCAAAGCTGGCGACAAGTAATTTCTGACTGAGAGAGAGTTTGAAGGAGACTTTGTTGTC | 1345 |
| QY | 1201 | GCTACATTCTTACCCGCTGAGTGCCTTACAGAGGGGATATCTGTATCAAACTTGCTTTAA | 1260 |
| Db | 1346 | GCTACATTCTTACCCGCTGAGTGCCTTACAGAGGGGATATCTGTATCAAACTTGCTTTAA | 1405 |
| RESULT 5 | | | |
| LOCUS | CQ412057 | 4065 bp | DNA |
| DEFINITION | Sequence 19128 from Patent WO0170979. | linear | PAT 23-JAN-2004 |
| ACCESSION | CQ412057 | | |
| VERSION | CQ412057.1 | GI:41319838 | |
| KEYWORDS | | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| AUTHORS | Lee, J. and Little, J. | | |
| TITLE | Genes, compositions, kits, and method for identification, assessment, prevention, and therapy of ovarian cancer | | |
| JOURNAL | Patent: WO 0170979-A 19128 27-SEP-2001; Millennium Pharmaceuticals, Inc. (US) | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..4065 | | |
| ORIGIN | 1..4065 | | |
| Query Match | 99.7%; Score 1256.8; DB 6; Length 4065; | | |
| Best Local Similarity | 99.8%; Pred. No. 0; | | |
| Matches 1258; Conservative | 0; Mismatches 2; Indels 0; Gaps 0; | | |
| QY | 1 | ATGGGCGGATCCGGCTGGCGCTGGGGGCGGAGACCGCGGCTTCTGAGAGGCTTGGG | 60 |
| Db | 268 | ATGGGCGGATCCGGCTGGCGCTGGGGGCGGAGACCGCGGCTTCTGAGAGGCTTGGG | 327 |
| QY | 61 | CGGCTGTGGCAGGTACAGAGCGCTGTGGGTAGCGGCTCTCCGCTCGGTGTATCGGTT | 120 |
| Db | 328 | CGGCTGTGGCAGGTACAGAGCGCTGTGGGTAGCGGCTCTCTCCGCTCGGTGTATCGGTT | 387 |
| QY | 121 | CGTGTCTGGGCAACCTGTGGCTGCCGCCCGCGCTCTCAAGACAGTTCTTGCCGACAGA | 180 |
| Db | 388 | CGTGTCTGGGCAACCTGTGGCTGCCGCCCGCGCTCTCAAGACAGTTCTTGCCGACAGA | 447 |
| QY | 181 | ACCAACCGGGGCGGGGCGCTCTGCGCGCGAGTAGTGTTCGCAAGAGAGGGCGGCGCTG | 240 |
| Db | 448 | ACCAACCGGGGCGGGGCGCTCTGCGCGCGAGTAGTGTTCGCAAGAGAGGGCGGCGCTG | 507 |
| QY | 241 | GAACAGTTGACAGGGTCACAGAAACATCGTGACTTTGTATGAGAGTGTTTACAATCACTTT | 300 |
| Db | 508 | GAACAGTTGACAGGGTCACAGAAACATCGTGACTTTGTATGAGAGTGTTTACAATCACTTT | 567 |
| QY | 301 | TCTCCAAATGTCGCATCAACGCTGTCTGTGCTTGAACCTCTGAGATGCAGTGTTCGAA | 360 |
| Db | 568 | TCTCCAAATGTCGCATCAACGCTGTCTGTGCTTGAACCTCTGAGATGCAGTGTTCGAA | 627 |
| QY | 361 | TTGCTCTTATATTCAGATCACCAAGGGTGTTCATGTGAGATGATACAGATTTGCGCCGA | 420 |
| Db | 628 | TTGCTCTTATATTCAGATCACCAAGGGTGTTCATGTGAGATGATACAGATTTGCGCCGA | 687 |
| QY | 421 | GATGTTTTGAGAGCCCTTGCTTTTCTTCATCAATGAGGGCTATGTCATGCGGACCTCAA | 480 |
| Db | 688 | GATGTTTTGAGAGCCCTTGCTTTTCTTCATCAATGAGGGCTATGTCATGCGGACCTCAA | 747 |
| QY | 481 | CCAGCTAACATTTGTGAGAGTGCACAGATGAATGTTTAAACTATTTGACTTTGGAATT | 540 |
| Db | 748 | CCAGCTAACATTTGTGAGAGTGCACAGATGAATGTTTAAACTATTTGACTTTGGAATT | 807 |

| | | | |
|------------|--|---|--------|
| QY | 541 | AGCTTCAAAGAGGCATATCGAATGTAAAGTATATTCAGACAGACGGGTATTCGGCTCCA | 600 |
| Db | 808 | AGCTTCAAAGAGGCATATCGAATGTAAAGTATATTCAGACAGACGGGTATTCGGCTCCA | 867 |
| QY | 601 | GAACGAGAAATTCGCAAAATTGCTTGCCAGAGCTGGCTCGAGAGTGAATACAGAAATGTACC | 660 |
| Db | 868 | GAACGAGAAATTCGCAAAATTGCTTGCCAGAGCTGGCTCGAGAGTGAATACAGAAATGTACC | 927 |
| QY | 661 | TCAGCTGTGATCTGTGTGAGGCTTACGAAATCATTTTATCTGGAATATGTTCTACAGAAATGAAA | 720 |
| Db | 928 | TCAGCTGTGATCTGTGTGAGGCTTACGAAATCATTTTATCTGGAATATGTTCTACAGAAATGAAA | 987 |
| QY | 721 | CTGAAACATACAGTCAGATCTCAGAGAAATGGAAGCAACAGTTCTGCTATTTATTTGATCAC | 780 |
| Db | 988 | CTGAAACATACAGTCAGATCTCAGAGAAATGGAAGCAACAGTTCTGCTATTTATTTGATCAC | 1047 |
| QY | 781 | ATATTTTCCAGTAAGAAGAGTGTGAATATCCGCCAATTCAGCTTATCACTTAAGAGACTTT | 840 |
| Db | 1048 | ATATTTTCCAGTAAGAAGAGTGTGAATATCCGCCAATTCAGCTTATCACTTAAGAGACTTT | 1107 |
| QY | 841 | ATCAAAAAGCATGCTTCAATGATATCCAAAGCAAGAAATTCCTGTGTGAATAGCATTTGTGC | 900 |
| Db | 1108 | ATCAAAAAGCATGCTTCAATGATATCCAAAGCAAGAAATTCCTGTGTGAATAGCATTTGTGC | 1167 |
| QY | 901 | AGCCCATTTCTTTAGCATTCCTCTTTGGCCCTCATTTGAAGATCTGTATCTTCCACT | 960 |
| Db | 1168 | AGCCCATTTCTTTAGCATTCCTCTTTGGCCCTCATTTGAAGATCTGTATCTTCCACT | 1227 |
| QY | 961 | CCAGTGTCAAGACTGTGAATGTCTGATGATGATTTATCTTGGAAATGAAGAGAAATAT | 1020 |
| Db | 1228 | CCAGTGTCAAGACTGTGAATGTCTGATGATGATTTATCTTGGAAATGAAGAGAAATAT | 1287 |
| QY | 1021 | GAAATGTGTGAAAGATGTGTAAGAAGAGAGTGTCAAAAATATGGAACAGATGTATCTCTA | 1080 |
| Db | 1288 | GAAATGTGTGAAAGATGTGTAAGAAGAGAGTGTCAAAAATATGGAACAGATGTATCTCTA | 1347 |
| QY | 1081 | CTTGTTCCAAAGAAAATTCCTGACAGAGGACAAGTCTTTGTGAGTATGCAAAATGCTGCT | 1140 |
| Db | 1348 | CTTGTTCCAAAGAAAATTCCTGACAGAGGACAAGTCTTTGTGAGTATGCAAAATGCTGCT | 1407 |
| QY | 1141 | GATTCCAAAGCTGGCGAGAAATTATCTGATCGAAGAGATGTTTGATGGAAGTTTGTGTG | 1200 |
| Db | 1408 | GATTCCAAAGCTGGCGAGAAATTATCTGATCGAAGAGATGTTTGATGGAAGTTTGTGTG | 1467 |
| QY | 1201 | GCTCATTTCTAACCCGCTGAGAGTCTTAAGAAGGGGATATCTGATCAAACTGCTTAA | 1260 |
| Db | 1468 | GCTCATTTCTAACCCGCTGAGAGTCTTAAGAAGGGGATATCTGATCAAACTGCTTAA | 1527 |
| RESULT 6 | | | |
| LOCUS | AX188296 | 4065 bp | linear |
| DEFINITION | Sequence 3991 from Patent WO0142467. | | |
| ACCESSION | AX188296 | | |
| VERSION | AX188296.1 GI:15139769 | | |
| KEYWORDS | . | | |
| SOURCE | Homo sapiens (human) | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| AUTHORS | 1 | | |
| TITLE | Schlegel, R., Deeds, J., Berger, A. and Zhao, X. | | |
| JOURNAL | Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer | | |
| FEATURES | Patent: WO 0142467-A 3991.14-UN-2001; Millennium Predictive Medicine, Inc. (US) | | |
| source | Location/Qualifiers | | |
| | 1..4065 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="unassigned DNA" | | |
| | /db_xref="taxon:9606" | | |
| ORIGIN | | | |

Query Match 99.7%; Score 1256.8; DB 6; Length 4065;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGCGGAGATCCGGCTGCGCTGCGGCGCGGAGCGCGCGTCTTCTGAGGCGCTTCGCG 60
DB ATGGCGGAGATCCGGCTGCGCTGCGGCGCGGAGCGCGCGTCTTCTGAGGCGCTTCGCG 327
QY 61 CGGCTGTGCAAGTACAGAGCGCTGTGGGTAGCGGCTCCCTCGCGCTCGGTATCGGGTT 120
DB 328 CGGCTGTGCAAGTACAGAGCGCTGTGGGTAGCGGCTCCCTCGCGCTCGGTATCGGGTT 387
QY 121 CGCTGTGCGGCAACCGTGGCTGCGCGCGCGCGCTCAAGCATCTTCTGCGCGCAGA 180
DB 388 CGCTGTGCGGCAACCGTGGCTGCGCGCGCGCGCTCAAGCATCTTCTGCGCGCAGA 447
QY 181 ACCACCGGCGCTGCGGCTCTGCGCGCGGATGTTTCCGCAAGAGAGGCGCGCTG 240
DB 448 ACCACCGGCGCTGCGGCTCTGCGCGCGGATGTTTCCGCAAGAGAGGCGCGCTG 507
QY 241 GAACAGTTCAGAGGTACAGAAACATCGTACTTGTATGAGAGTGTTCATCCACTTT 300
DB 508 GAACAGTTCAGAGGTACAGAAACATCGTACTTGTATGAGAGTGTTCATCCACTTT 567
QY 301 TCTCCAAATGTGCGCATCAGCTGTCTGTGCTTGAATCCTCGGATGTCAGTGTTCGGA 360
DB 568 TCTCCAAATGTGCGCATCAGCTGTCTGTGCTTGAATCCTCGGATGTCAGTGTTCGGA 627
QY 361 TTGCTCTTATATTCAGATCAGAGGCTGTTCATGTGATGATACAGATTCGCGCGA 420
DB 628 TTGCTCTTATATTCAGATCAGAGGCTGTTCATGTGATGATGATGATGATGATGATG 687
QY 421 GATGTTTGGAGGCGCTGCTTCTTCTCATATGAGGCGATGTCATGCGGACCTCAAA 480
DB 688 GATGTTTGGAGGCGCTGCTTCTTCTCATATGAGGCGATGTCATGCGGACCTCAAA 747
QY 481 CCACGTATATATTTGAGTGCAGAGATGAAATGTTTAACTCATTTGAGACTT 540
DB 748 CCACGTATATATTTGAGTGCAGAGATGAAATGTTTAACTCATTTGAGACTT 807
QY 541 AGCTTCAAAAGAGGCAATCAGATGTAAATATTTCAAGACAGCGGATTCGGGCTCCA 600
DB 808 AGCTTCAAAAGAGGCAATCAGATGTAAATATTTCAAGACAGCGGATTCGGGCTCCA 867
QY 601 GAAGCAGATTTGCAAAATGCTTGCGCGAGGCTGCGCGAGTATACAGATGTACC 660
DB 868 GAAGCAGATTTGCAAAATGCTTGCGCGAGGCTGCGCGAGTATACAGATGTACC 927
QY 661 TCAGCTGTGATCTGTGAGGCTTGAAGATCAATTTTCTGAAAATGTTCTCAGGAATGAAA 720
DB 928 TCAGCTGTGATCTGTGAGGCTTGAAGATCAATTTTCTGAAAATGTTCTCAGGAATGAAA 987
QY 721 CTGAAAATATCAGATCTCAGAAATGAGAGGCAAAACAGTCTGTATATTTGATCAC 780
DB 988 CTGAAAATATCAGATCTCAGAAATGAGAGGCAAAACAGTCTGTATATTTGATCAC 1047
QY 781 ATATTGGCAGTAAAGAGTGTGAATGCGCAATTTCAAGCTTATCACTTAAGAGACTT 840
DB 1048 ATATTGGCAGTAAAGAGTGTGAATGCGCAATTTCAAGCTTATCACTTAAGAGACTT 1107
QY 841 ATCAAAAGATGCTTCAATGATTCAGAGCAAGAAATTCCTGCTGAATGCGATTTGCG 900
DB 1108 ATCAAAAGATGCTTCAATGATTCAGAGCAAGAAATTCCTGCTGAATGCGATTTGCG 1167
QY 901 AGCCCATTTCTTATGCAATCTTTTGGCCCTCATATTTGAAGATCTGCTATCCCACT 960
DB 1168 AGCCCATTTCTTATGCAATCTTTTGGCCCTCATATTTGAAGATCTGCTATCCCACT 1227
QY 961 CCAGTCTTAAGATGCTGAGATGTGCTGAGATGATTTATCTTGGAGATGAAGAGAAAT 1020
DB 1228 CCAGTCTTAAGATGCTGAGATGTGCTGAGATGATTTATCTTGGAGATGAAGAGAAAT 1287
```

```
QY 1021 GAAGATGTTTGAAGATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTA 1080
DB 1288 GAAGATGTTTGAAGATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTA 1347
QY 1081 CTTGTTCCAAAGAAATTCCTGGCAGAGACAAAGCTTTTGTGATGATGCAAAATGCTGT 1140
DB 1348 CTTGTTCCAAAGAAATTCCTGGCAGAGACAAAGCTTTTGTGATGATGCAAAATGCTGT 1407
QY 1141 GATTCCAAAGCTGCGCAGAAATTTACTGACTGGAAGATGTTTGTATGAGAAATGTTGTG 1200
DB 1408 GATTCCAAAGCTGCGCAGAAATTTACTGACTGGAAGATGTTTGTATGAGAAATGTTGTG 1467
QY 1201 GCTAATTTCTACCCGCTGAGTCTTCAAGAGGAGATCTGTATCAAACTTGTCTTAA 1260
DB 1468 GCTAATTTCTACCCGCTGAGTCTTCAAGAGGAGATCTGTATCAAACTTGTCTTAA 1527
```

RESULT 7
AX16543 1260 bp DNA linear PAT 22-JUN-2001
LOCUS
DEFINITION
Sequence 34 from Patent WO0138503.
ACCESSION
AX16543
VERSION
AX16543.1 GI:14546888

KEYWORDS

SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS
Ploewman, G.D., Whyte, D., Manning, G.S., Sudarshan, S.S., Martinez, R.,
Flanagan, P., and Clardy, D.S.
TITLE
Novel human protein kinases and protein kinase-like enzymes
JOURNAL
Patent: WO 0138503-A 34 31-MAY-2001;
Sugen, Inc. (US)

FEATURES

source
1..1260
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.6%; Score 1255.2; DB 6; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 ATGGCGGAGATCCGGCTGCGCTGCGGCGCGGAGCGCGCGTCTTCTGAGGCGCTTCGCG 60
DB 1 ATGGCGGAGATCCGGCTGCGCTGCGGCGCGGAGCGCGCGTCTTCTGAGGCGCTTCGCG 60
QY 61 CGGCTGTGCAAGTACAGAGCGCTGTGGGTAGCGGCTCCCTCGCGCTCGGTATCGGGTT 120
DB 61 CGGCTGTGCAAGTACAGAGCGCTGTGGGTAGCGGCTCCCTCGCGCTCGGTATCGGGTT 120
QY 121 CGCTGTGCGGCAACCGTGGCTGCGCGCGCGCGCGCTCAAGCATCTTTCGCGCAGA 180
DB 121 CGCTGTGCGGCAACCGTGGCTGCGCGCGCGCGCGCTCAAGCATCTTTCGCGCAGA 180
QY 181 ACCACCGGCGCTGCGGCTCTGCGCGCGGATGTTTCCGAAAGAGAGGCGCGCTG 240
DB 181 ACCACCGGCGCTGCGGCTCTGCGCGCGGATGTTTCCGAAAGAGAGGCGCGCTG 240
QY 241 GAACAGTTCAGAGGTACAGAAACATCGTACTTGTATGAGAGTGTTCATCCACTTT 300
DB 241 GAACAGTTCAGAGGTACAGAAACATCGTACTTGTATGAGAGTGTTCATCCACTTT 300
QY 301 TCTCCAAATGTGCGCATCAGCTGTCTGTGCTTGAATCCTCGGATGTCAGTGTTCGGA 360
DB 301 TCTCCAAATGTGCGCATCAGCTGTCTGTGCTTGAATCCTCGGATGTCAGTGTTCGGA 360
QY 361 TTGCTCTTATATTCAGATCAGAGGCTGTTCATGTGATGATACAGATTCGCGCGA 420
DB 361 TTGCTCTTATATTCAGATCAGAGGCTGTTCATGTGATGATGATGATGATGATGATG 420
```

QY 421 GATGTTTGAGAGCCCTGCTTTCTTCATCATGAGGCGTATGTCAGTGGAGCTCAAA 480
 DB 421 GATGTTTGAGAGCCCTGCTTTCTTCATCATGAGGCGTATGTCAGTGGAGCTCAAA 480
 QY 481 CCAAGTAAATATGAGAGTGAAGTAAATGTTTAACTCAATGACTTTGACTT 540
 DB 481 CCAAGTAAATATGAGAGTGAAGTAAATGTTTAACTCAATGACTTTGACTT 540
 QY 541 AGCTTCAAGAGGAGCATAGATGTAATGTAATTCAGACAGACGGGATTCGGCTCCA 600
 DB 541 AGCTTCAAGAGGAGCATAGATGTAATGTAATTCAGACAGACGGGATTCGGCTCCA 600
 QY 601 GAAGCAGAAATGCAAAATTCCTGGCCAGAGCTGCGCTGAGAGTACAGATATACC 660
 DB 601 GAAGCAGAAATGCAAAATTCCTGGCCAGAGCTGCGCTGAGAGTACAGATATACC 660
 QY 661 TCAGCTGTTGATCTGAGAGCTGAGATCAATTTTACTGGAATGTTCTCAGGAATGAAA 720
 DB 661 TCAGCTGTTGATCTGAGAGCTGAGATCAATTTTACTGGAATGTTCTCAGGAATGAAA 720
 QY 721 CTGAAACATACAGTCAATCTCAGAAATGGAAGGCAAAACAGTTCTGCTATATATGATCAC 780
 DB 721 CTGAAACATACAGTCAATCTCAGAAATGGAAGGCAAAACAGTTCTGCTATATATGATCAC 780
 QY 781 ATATTGCGCAGTAAAGCAGTGAATGCGCAATTCAGACCTATACCTAAGAGACTT 840
 DB 781 ATATTGCGCAGTAAAGCAGTGAATGCGCAATTCAGACCTATACCTAAGAGACTT 840
 QY 841 ATCAAAAGCATGCTTCATGATGATCCAGAGAAATTCCTGCTGAAATGGCAATGTGC 900
 DB 841 ATCAAAAGCATGCTTCATGATGATCCAGAGAAATTCCTGCTGAAATGGCAATGTGC 900
 QY 901 AGCCCATTTCTTGAATCTCTTTGCGCCCTCATATTTGAAGATCTGCTCATGCTTCCACT 960
 DB 901 AGCCCATTTCTTGAATCTCTTTGCGCCCTCATATTTGAAGATCTGCTCATGCTTCCACT 960
 QY 961 CCAGTCTAAGACTGCTGAATGCTGATGATGATTTATCTGGAATGAAGAGAAATAT 1020
 DB 961 CCAGTCTAAGACTGCTGAATGCTGATGATGATTTATCTGGAATGAAGAGAAATAT 1020
 QY 1021 GAAGATGTTGTAGAGATGTAAGAGAGAGTGTCAAAATATATGACCAATGATCTCTA 1080
 DB 1021 GAAGATGTTGTAGAGATGTAAGAGAGAGTGTCAAAATATATGACCAATGATCTCTA 1080
 QY 1081 CTTGTTCCAAAGAAATCTGCGACAGAGCAAGTCTTTGTTGAATGAATGCTGCT 1140
 DB 1081 CTTGTTCCAAAGAAATCTGCGACAGAGCAAGTCTTTGTTGAATGAATGCTGCT 1140
 QY 1141 GATTCCAAAGCTGCGAGAAATTAATGACTGGAAGAGTGTGGAAGAAAGTTGTTG 1200
 DB 1141 GATTCCAAAGCTGCGAGAAATTAATGACTGGAAGAGTGTGGAAGAAAGTTGTTG 1200
 QY 1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGATCAAACTTGTCTTAA 1260
 DB 1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGATCAAACTTGTCTTAA 1260

RESULT 8
 AX680136 1824 bp DNA linear PAT 29-MAR-2003
 LOCUS Sequence 21 from Patent WO0181555.
 DEFINITION AX680136
 ACCESSION AX680136
 VERSION AX680136.1 GI:29369916
 KEYWORDS
 ORGANISM Homo sapiens (human)
 SOURCE Homo sapiens
 BUKARYOZA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 1 Gururajan, R., Paterson, C., Recipon, S. A., Tribouley, C. M.,
 Hafalla, A., Khan, F. Y., Yue, H., Au-Young, J., Bandman, O.,
 Baughin, M. R., Borowsky, M. D., Burford, N., Burdill, J. D., Elliott, V. S.,
 Gandhi, A. R., Kearney, L., Lal, P., Lu, D. A., Lu, Y., Tang, Y. T.,

Tingler, K. A., Griffin, J. A., Hillman, J. J., Marcus, G. A., Nguyen, D. B.,
 Policky, J., Ramkumar, J., Thornton, M., Walla, N. K. and Walsh, R. T.
 Human Kinases
 Patent: WO 0181555-A 21 01-NOV-2001;
 Incyte Genomics, Inc. (US)
 Location/Qualifiers
 1..1824
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 3107952CB1"

ORIGIN

Query Match 99.6%; Score 1255.2; DB 6; Length 1824;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATGGCGGATCCGGCTGCGCTGCGGCGGAGCCGCGTTTCTGAGGCTTCGCG 60
 DB 169 ATGGCGGATCCGGCTGCGGCTGCGGCGGAGCCGCGTTTCTGAGGCTTCGCG 228
 QY 61 CCGCTGTCAGAGTACAGAGCTGCTGAGAGCGCTCTCGCTCGGTATGCGGT 120
 DB 229 CCGCTGTCAGAGTACAGAGCTGCTGAGAGCGCTCTCGCTCGGTATGCGGT 288
 QY 121 CGCTGTCGCGACCCCTGCGCTGCGCCCGCGGCTCAAGCAGTTCTGCGCAGGA 180
 DB 289 CGCTGTCGCGACCCCTGCGCTGCGCCCGCGGCTCAAGCAGTTCTGCGCAGGA 348
 QY 181 ACCACCGGCGCTGCGGCTGCTGCGCGAGTATGTTTCCGAAAGAGGCGCGCTG 240
 DB 349 ACCACCGGCGCTGCGGCTGCTGCGCGAGTATGTTTCCGAAAGAGGCGCGCTG 408
 QY 241 GAACGTTGCGAGGTACAGAAACATCTGACTTTGATGAGTGTTCATCAATCTT 300
 DB 409 GAACGTTGCGAGGTACAGAAACATCTGACTTTGATGAGTGTTCATCAATCTT 468
 QY 301 TCTCCAAATGTCGATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 DB 469 TCTCCAAATGTCGATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
 QY 361 TTGCTCTTATATTCAGTACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 529 TTGCTCTTATATTCAGTACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 588
 QY 421 GATGTTTGAAGCCCTGCTTTCTTATCATGAGGCTATGTCAGAGGAGCTCAAA 480
 DB 589 GATGTTTGAAGCCCTGCTTTCTTATCATGAGGCTATGTCAGAGGAGCTCAAA 648
 QY 481 CCAAGTAAATATGAGAGTGAAGTAAATGTTTAACTCAATGACTTTGACTT 540
 DB 649 CCAAGTAAATATGAGAGTGAAGTAAATGTTTAACTCAATGACTTTGACTT 708
 QY 541 AGCTTCAAGAGGAGCATAGAGTGAAGTAAATGTTTAACTCAATGACTTTGACTT 600
 DB 709 AGCTTCAAGAGGAGCATAGAGTGAAGTAAATGTTTAACTCAATGACTTTGACTT 768
 QY 601 GAAGCAGAAATGCAAAATTCCTGGCCAGAGCTGCGCTGAGAGTACAGATATACC 660
 DB 769 GAAGCAGAAATGCAAAATTCCTGGCCAGAGCTGCGCTGAGAGTACAGATATACC 828
 QY 661 TCAGCTGTTGATCTGAGAGCTGAGATCAATTTTACTGGAATGTTCTCAGGAATGAAA 720
 DB 829 TCAGCTGTTGATCTGAGAGCTGAGATCAATTTTACTGGAATGTTCTCAGGAATGAAA 888
 QY 721 CTGAAACATACAGTCAATCTCAGAAATGGAAGGCAAAACAGTTCTGCTATATATGATCAC 780
 DB 889 CTGAAACATACAGTCAATCTCAGAAATGGAAGGCAAAACAGTTCTGCTATATATGATCAC 948
 QY 781 ATATTGCGCAGTAAAGCAGTGAATGCGCAATTCAGACCTATACCTAAGAGACTT 840
 DB 949 ATATTGCGCAGTAAAGCAGTGAATGCGCAATTCAGACCTATACCTAAGAGACTT 1008

QY 841 ATCAAAAGCATGCTTCATGATGATCCAGAGAGAATTCCTGCGTGAATGGCATTTGTC 900
 DB 1009 ATCAAAAGCATGCTTCATGATGATCCAGAGAGAATTCCTGCGTGAATGGCATTTGTC 1068
 QY 901 AGCCCATTTCTTTCAGCATTTCTTTTGGCCCTCATATTGGAAGATTTGGTCACTGCTCCACT 960
 DB 1069 AGCCCATTTCTTTCAGCATTTCTTTTGGCCCTCATATTGGAAGATTTGGTCACTGCTCCACT 1128
 QY 961 CCAGTCTAAGACGTCTGATGATGCTGATGATGATTTATCTTTGGGAATGAAGAGAAATAT 1020
 DB 1129 CCAGTCTAAGACGTCTGATGATGCTGATGATGATTTATCTTTGGGAATGAAGAGAAATAT 1188
 QY 1021 GAAGATGTTGTAGAGATGTAAGAGAGAGATGTCACAAAATATGACCAAGTGTATCTCTA 1080
 DB 1189 GAAGATGTTGTAGAGATGTAAGAGAGAGATGTCACAAAATATGACCAAGTGTATCTCTA 1248
 QY 1081 CTTGTTCCAAAGGAAATCTCTGCGACAGCAAGTCTTTTGTGATGATGATGATGATGATGATGAT 1140
 DB 1249 CTTGTTCCAAAGGAAATCTCTGCGACAGCAAGTCTTTTGTGATGATGATGATGATGATGATGAT 1308
 QY 1141 GATTCCAAAGCTGCGAGAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1309 GATTCCAAAGCTGCGAGAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1368
 QY 1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTTGCTTTAA 1260
 DB 1369 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTTGCTTTAA 1428
 RESULT 9
 LOCUS BC058732
 DEFINITION Mus musculus kinase interacting with leukemia-associated gene (setachman), mRNA (CDNA clone MGC:64770 IMAGE:6414877), complete cds.
 ACCESSION BC058732 2041 bp mRNA linear ROD 30-JUN-2004
 VERSION BC058732.1 GI:37194892
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2041)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleorn, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ueblich, T.B., Toshiyuki, S., Carinici, P., Prange, C., Raha, S.S., Loggiano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, D.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skelske, U., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 TITLE human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 12477932
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (24-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT
 Contact: MGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Ronald, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Smir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnlnl.gov>
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754447.
 Location/Qualifiers
 1..2041
 /organism="Mus musculus"
 /mol_type="mRNA"
 /accession="U00000"
 /db_xref="taxon:10090"
 /clone="MGC:64770 IMAGE:6414877"
 /tissue_type="Brain, pool of enriched mouse brain tissues from 13.5, 14.5, 16.5 and 17.5 dpc"
 /clone_lib="NIH BMAP_FRO"
 /lab_host="DH10B"
 /map="vector: pYX-ASC"
 1..2041
 /gene="Kist"
 /note="synonym: KIS"
 /db_xref="LOCUSID:16589"
 /db_xref="MGI:1341908"
 154..1413
 /gene="Kist"
 /codon_start=1
 /product="kinase interacting with leukemia-associated gene (setachman)"
 /protein_id="AAH58732.1"
 /db_xref="GI:37194893"
 /db_xref="LOCUSID:16589"
 /db_xref="MGI:1341908"
 /translation="MAGSGCAGWAEPPFLFAFRLMVOVSLGSSASVYRACCG TPGSPGALKOPFLPGTGAASABYGFRRERALEOGLRNVLTLYGVFTIFSP NVSPRCILBLIDVSESLLYSSHSGSMMMOICARDALEALFLHREGVADLK PRNITMSANERFKLIDGLSPKESNOVXYIOTGVAPAPAEONCIAAGLOSTDE CTSAVDLMSLGLILEMFSGMKDKRTVRSQEKANSSATIDHIFASKAVNADLPAH LRLDKSLMHDPGRIRIPAEMLCSPPFSIPAPHIDLVMLPTVPLRLNLVLDYDL ENEDYEVNVEVEKRCORGYVSLVLPKKNPGGVFVEYVANAAGSKAOKLTLR MFDGFVATFVPLSAVRYGLYQLL"
 ORIGIN
 Query Match 88.6%; Score 1116; DB 10; Length 2041;
 Best Local Similarity 92.9%; Pred. No. 1.5e-278;
 Matches 1170; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
 QY 1 ATGGGCGGATCCGGGCTGCGGCGCGGAGCGCGCGGCTTTCTGAGAGCGCTTCCGG 60
 DB 154 ATGGGCGGATCCGGGCTGCGGCGCGGAGCGCGCGGCTTTCTGAGAGCGCTTCCGG 213
 QY 61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTCGGTGTATCGGGT 120
 DB 214 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTCGGTGTATCGGGT 273
 QY 121 CGCTGTGGGCAACCTGCGCTGCGCCCGCGCGCGCTTCAAGCATTTCTGCGCGCAGGA 180
 DB 274 CGCTGTGGGCAACCTGCGCTGCGCCCGCGCGCGCTTCAAGCATTTCTGCGCGCAGGA 333
 QY 181 ACCACCGGGGCTGCGGCTCTGCGCGCGGTATGTTCCGCAAGAGAGGCGGCGCTG 240

```

Db      334 ACCACGGGGCTGGGCTTCGGCCGGAGATGTTTCGCAAGAGAGGGCGGCTG 393
Qy      241 GAACAGTTGAGGGGTCAAGAAACATCGTACTTTGATGAGAGTGTTCACATCACTT 300
Db      394 GAGCAGTTGAGGGGTCAAGAAACATCGTACTTTGATGAGAGTGTTCACATCACTT 453
Qy      301 TCTCCAAATGTCGCATCAGCGTGTCTGTTGAACTCCGTGAGTGCAGTGTTCGGA 360
Db      454 TCTCCAAATGTCGCATCAGCGTGTCTGTTGAACTCCGTGAGTGCAGTGTTCGGA 513
Qy      361 TTGCTCTATATTCAGTCAACGAGGTGTTCAATGATGATACAGCATTCGCCCA 420
Db      514 TTGCTCTATATTCAGTCAACGAGGTGTTCAATGATGATACAGCATTCGCCCA 573
Qy      421 GATGTTTGGAGGCGCTGCTTTCTTCATCATAGAGGCTATGCTCATGCGACTCA 480
Db      574 GATGTTTGGAGGCGCTGCTTTCTTCATCATAGAGGCTATGCTCATGCGACTCA 633
Qy      481 CCAAGTAAATATTTGAGTGCAGAGAAATGATTTTAACTCAATTCATTTGACTT 540
Db      634 CCAAGTAAATATTTGAGTGCAGAGAAATGATTTTAACTCAATTCATTTGACTT 693
Qy      541 AGCTTCAAGAGAGGCATCAGATGTAAGTATTTCAAGACAGGCTATCGGCTCCA 600
Db      694 AGCTTCAAGAGAGGCATCAGATGTAAGTATTTCAAGACAGGCTATCGGCTCCA 753
Qy      601 GAAGCAGAAATGCAAAATGCTTGGCCAGGCTGGCTGCAGAGTATCAGAAATGAC 660
Db      754 GAAGCAGAAATGCAAAATGCTTGGCCAGGCTGGCTGCAGAGTATCAGAAATGAC 813
Qy      661 TCAGCTGTGATCTGAGAGCTTGAAGATCATTTTCTGGAATGTTCTCAGAAATGA 720
Db      814 TCAGCTGTGATCTGAGAGCTTGAAGATCATTTTCTGGAATGTTCTCAGAAATGA 873
Qy      721 CTGAAACATACAGTCAAGATCTCAGAAATGAGGCAACAGTTCTGCTATTTATGAC 780
Db      874 CTGAAACATACAGTCAAGATCTCAGAAATGAGGCAACAGTTCTGCTATTTATGAC 933
Qy      781 ATATTTGCAAGTAAAGAGTGAATGCGCAATTCAGGCTATGCACTTAAGAGCTT 840
Db      934 ATATTTGCAAGTAAAGAGTGAATGCGCAATTCAGGCTATGCACTTAAGAGCTT 993
Qy      841 ATCAAAAGATGCTTATGATGATTCAGAGCAAGAAATTCCTGCTGAATGCAATGTC 900
Db      994 ATCAAAAGATGCTTATGATGATTCAGAGCAAGAAATTCCTGCTGAATGCAATGTC 1053
Qy      901 AGCCCATCTTGAAGATCTCTTTGCTTGCCTCATATTTGAAGATCTGATCTGCCACT 960
Db      1054 AGCCCATCTTGAAGATCTCTTTGCTTGCCTCATATTTGAAGATCTGATCTGCCACT 1113
Qy      961 CCAGTCTAAAGATGCTGTAATGCTGATGATGATTTATCTTGGAAATGAAGAGAAAT 1020
Db      1114 CCAGTCTAAAGATGCTGTAATGCTGATGATGATTTATCTTGGAAATGAAGAGAAAT 1173
Qy      1021 GAAGATGTTGTAAGATGTAAGAGAGATGCAAAAATATGACCAAGTGTATCTTA 1080
Db      1174 GAAGATGTTGTAAGATGTAAGAGAGATGCAAAAATATGACCAAGTGTATCTTA 1233
Qy      1081 GTTGTTCAAAGAAATCTTGCGCAGAGCAAGTCTTGTGTAGATGCAATGCTGTGT 1140
Db      1234 GTTGTTCAAAGAAATCTTGCGCAGAGCAAGTCTTGTGTAGATGCAATGCTGTGT 1293
Qy      1141 GATTCCAAAGCTGCGGAGAAATTAATGATCTGAAAGATGTTGATGAGAAATGTTGTG 1200
Db      1294 GATTCCAAAGCTGCGGAGAAATTAATGATCTGAAAGATGTTGATGAGAAATGTTGTG 1353
Qy      1201 GCTACATTTACCGCTGAGTGCCTCAAGAGGGGATATCTGTATCAAACTTGTCTTA 1260
Db      1354 GCTACATTTACCGCTGAGTGCCTCAAGAGGGGATATCTGTATCAAACTTGTCTTA 1413

```

RESULT 10
AC129141_1/c

```

WPCOMMENT
Sequence split into 4 fragments LOCUS AC129141 Accession AC129141
Fragment Name      Begin      End
AC129141_0         1          110000
AC129141_1        100001     210000
AC129141_2        200001     310000
AC129141_3        300001     363206
Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH2)

Query Match      88.3%; Score 1112.8; DB 2; Length 110000;
Best Local Similarity 92.7%; Pred. No. 2.3e-277;
Matches 1168; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy      1 ATGCGCGGATCCGGGCTGCGCTCGGAGCGGAGCGCGGCTTCTGAGAGCTTCCGG 60
Db      70316 ATGCGCGGATCCGGGCTGCGCTCGGAGCGGAGCGCGGCTTCTGAGAGCTTCCGG 70257
Qy      61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTACCGGCTCTCCGCTCGGTGTATCGGTT 120
Db      70256 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTACCGGCTCTCCGCTCGGTGTATCGGTT 70197
Qy      121 CGCTGTGGGCAACCTGTGCTGCCCCCGGCGCTTCAGACATTTCTTCCGCGCAGGA 180
Db      70196 CGCTGTGGGCAACCTGTGCTGCCCCCGGCGCTTCAGACATTTCTTCCGCGCAGGA 70137
Qy      181 ACCACGGGGCTGGGCGCTCTGCGCGCGAGTATGTTTCCGCAAGAGAGGGCGGCTG 240
Db      70136 ACCACGGGGCTGGGCGCTCTGCGCGCGAGTATGTTTCCGCAAGAGAGGGCGGCTG 70077
Qy      241 GAACAGTTGCAAGGTCACAGAAACATCGTACTTTGATGAGAGTGTTCACATCACTT 300
Db      70076 GAACAGTTGCAAGGTCACAGAAACATCGTACTTTGATGAGAGTGTTCACATCACTT 70017
Qy      301 TCTCCAAATGTCGCATCAGCGTGTCTGTTGTAATCTCTGATGTCAAGTGTTCGGA 360
Db      70016 TCTCCAAATGTCGCATCAGCGTGTCTGTTGTAATCTCTGATGTCAAGTGTTCGGA 69957
Qy      361 TTGCTCTATATTCAGTCAACGAGTGTTCAGTGAAGATACAGATTTGGGCGCGGA 420
Db      69956 TTGCTCTATATTCAGTCAACGAGTGTTCAGTGAAGATACAGATTTGGGCGCGGA 69897
Qy      421 GATGTTTGGAGGCGCTTCTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 480
Db      69896 GATGTTTGGAGGCGCTTCTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 69837
Qy      481 CCAAGTAAATATTTGATGATGTCAGAGAAATGATTTTAACTCATGACTTTGACTT 540
Db      69836 CCAAGTAAATATTTGATGATGTCAGAGAAATGATTTTAACTCATGACTTTGACTT 69777
Qy      541 AGCTTCAAGAGAGCAATCAGATGTAAGTATTCAGACAGAGGCTATCGGCTCCA 600
Db      69776 AGCTTCAAGAGAGCAATCAGATGTAAGTATTCAGACAGAGGCTATCGGCTCCA 69717
Qy      601 GAAGCAGAAATGCAAAATGCTTGGCCAGGCTGCGTGCAGAGTATCAGAAATGTAAC 660
Db      69716 GAAGCAGAAATGCAAAATGCTTGGCCAGGCTGCGTGCAGAGTATCAGAAATGTAAC 69657
Qy      661 TCAGCTGTGATCTGTGAGAGCTTGAAGATCAATTTACTGAAATGTTCTCAGAAATGAA 720
Db      69656 TCAGCTGTGATCTGTGAGAGCTTGAAGATCAATTTACTGAAATGTTCTCAGAAATGAA 69597
Qy      721 CTGAAACATACAGTACAGATCTCAGAAATGAGAGCAACAGTTCTGCTATTTATGATAC 780
Db      69596 CTGAAACATACAGTACAGATCTCAGAAATGAGAGCAACAGTTCTGCTATTTATGATAC 69537
Qy      781 ATATTTGCAAGTAAAGAGTGTGTAATCGCAATTCAGGCTTACCTTAAGAGACCTT 840
Db      69536 ATATTTGCAAGTAAAGAGTGTGTAATCGCAATTCAGGCTTACCTTAAGAGACCTT 69477
Qy      841 ATCAAAAGATGCTTATGATGATTCAGAGCAAGAAATCTGCTGAATGCAATTTGTC 900
Db      69476 ATCAAAAGATGCTTATGATGATTCAGAGCAAGAAATCTGCTGAATGCAATTTGTC 69417

```

| | | | |
|------------|---|--|-----------------|
| QY | 901 | AGCCATCTCTTACCATCTCCCTTTGGCCCTCATATGTGAAGATCTGTCATCTCCCACT | 960 |
| Db | 69116 | AGCCATCTCTTACCATCTCCCTTTGGCCCTCATATGTGAAGATCTGTCATCTCCCACT | 6935 |
| QY | 961 | CCAGTGTCTAAGACTGCTGGAATGCTGCGATGATGATTAATCTTGGGAATGAAGAGAAATAT | 1020 |
| Db | 69356 | CCAGTGTCTAAGACTGCTGGAATGCTGCGATGATGATTAATCTTGGGAATGAAGAGAAATAT | 6929 |
| QY | 1021 | GAAGATGTGTGAAGAAATGTAAAAAGAGAGTGCATAAAATATGACACAGTGTCTACTA | 1080 |
| Db | 69296 | GAAGATGTGTGAAGAAATGTAAAAAGAGAGTGCATAAAATATGACACAGTGTCTACTG | 6923 |
| QY | 1081 | CTTGTTCCAAAGAAAAATCTCGCAGAGAGACAATCTTTGTTGATATGCAAAATGCTGT | 1140 |
| Db | 69236 | CTTGTTCCAAAGAAAAATCTCGCAGAGAGACAATCTTTGTTGATATGCAAAATGCTGTG | 6917 |
| QY | 1141 | GATTCCAAAGCTGGCGCAGAAATTAATCTGATCGAAGAGATGTTGATGGGAATTTGTTGTG | 1200 |
| Db | 69176 | GATTCCAAAGCTGGCGCAGAAATTAATCTGATCGAAGAGATGTTGATGGGAATTTGTTGTG | 6911 |
| QY | 1201 | GCTACATTTCTACCCGCTGAGAGCCTACAAAGAGGGAATATCTGTATCAAACTTGTCTTAA | 1260 |
| Db | 69116 | GCTACATTTCTACCCGCTGAGAGCCTACAAAGAGGGAATATCTTTATCAAACTTGTCTTAA | 6905 |
| RESULT 11 | | | |
| LOCUS | RNKISRNA | 1633 bp | linear |
| DEFINITION | R. norvegicus mRNA for Kis protein. | | ROD 08-JUL-2002 |
| ACCESSION | X98374 | | |
| VERSION | X98374.1 | GI:1403531 | |
| KEYWORDS | Kis gene. | | |
| SOURCE | Rattus norvegicus (Norway rat) | | |
| ORGANISM | Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. | | |
| REFERENCE | 1 | | |
| AUTHORS | Maucuer, A., Ozon, S., Manceau, V., Gavec, O., Lawler, S., Curmi, P. and Sobel, A. | | |
| TITLE | Kis is a protein kinase with an RNA recognition motif | | |
| JOURNAL | J. Biol. Chem. 272 (37), 23151-23156 (1997) | | |
| MEDLINE | 97435279 | | |
| PUBMED | 9287318 | | |
| REFERENCE | 2 (bases 1 to 1633) | | |
| AUTHORS | Maucuer, A. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer Moulin, 75005 Paris, FRANCE | | |
| FEATURES | | | |
| source | Location/Qualifiers | | |
| | 1..1633 | | |
| | /organism="Rattus norvegicus" | | |
| | /mol_type="mRNA" | | |
| | /db_xref="taxon:10116" | | |
| | /tissue_type="brain" | | |
| | /clone_lib="7" | | |
| | /dev_stage="neonate" | | |
| | 1..1633 | | |
| | /gene="Kis" | | |
| | 39..1298 | | |
| | /gene="Kis" | | |
| | /codon_start=1 | | |
| | /protein_id="CAA67021.1" | | |
| | /db_xref="GI:1403532" | | |
| | /db_xref="GOA:O63285" | | |
| | /db_xref="Swiss-Prot:O63285" | | |
| | /translation="MAGSGCAMGAEPPEFLFAGRLWQVOSRLGSGSSAYVRCCGCTPGSPBAIKQFLPRTGTGAASAAYGFRERARLQSHRIVLIVLVGVTHTSPBNVPERCLILBLDVSYSELIVYSHQCSMMIHCARDVLEALPLHBSYVADLKPRTNLWSAENECFKLIDFGLSFKGNQDVKTIQDTGRAPAEIQNCIAQGLSDTEPTSLAVDLMSLGIITLLEWFSGMKLKHTVRSQEWKNSAIIIDHIFASRAVVAALPAVHLRDLIKEMLHDDPSRRIPAEWALCSPPSIPFAHIDLVWLPTPVRLINLVDDVILLENBEYVDVDEVEECQKYPVSVLVPKENPERGVFVYNAAGDSKAQOKLITGR | | |
| gene | | | |
| CDS | | | |

| ORIGIN | MFDGKEVATFYPDLASVKRXYLTULL" | | | | | | | | | |
|-----------------------|-----------------------------|---|-----------|--------------|--|--|--|--|--|--|
| Query Match | 87.9% | Score 1108; | DB 10; | Length 1633; | | | | | | |
| Best Local Similarity | 92.5% | Pred. No. 1.7e-276; | | | | | | | | |
| Matches 1165; | Conservative 0; | Mismatches 95; | Indels 0; | Gaps 0; | | | | | | |
| Qy | 1 | ATGGCGGAGATCCGAGCTCGCTGGCGGGGCGCGAGCCGCGGTTTTCTGAGAGCCCTTCGGG | 60 | | | | | | | |
| Db | 39 | ATGGCCGGGCTCCGGCTCGCGTGGGGGCGGAGCCGCGGGCTTCTGGAGGCTTCGGG | 98 | | | | | | | |
| Qy | 61 | CGGCTGTGCAGGATACAGAGCCCTCTGGGTAAGCGGCTCTCCGCTCGGTGTATCGGCTT | 120 | | | | | | | |
| Db | 99 | CGGCTGTGCAGGATACAGAGCCGCTTGGGACGCGGCTCTCCGCGCTCGGTGTATCGGCTT | 158 | | | | | | | |
| Qy | 121 | CGCGTGCGGGCAACCCGTGGCTGGCCCCCGGGGCGCTCAAGCAGTTCTTGCGCCGACAGA | 180 | | | | | | | |
| Db | 159 | CGCGTGCGGGCACTCCAGGCTGGCCCCCGGGGCGCTCAAGCAGTTCTTGCTTCGGGA | 218 | | | | | | | |
| Qy | 181 | ACCAACCGGGGCTGGCGGCTCTGCGCGCGAGTATGGTTTCGCAAGAGAGGGCGGCGCTG | 240 | | | | | | | |
| Db | 219 | ACCAACCGGGGCTGGCGGCTCTGGCGGCGAGTATGGTTTCGCAAGAGAGGGCGGCGCTG | 278 | | | | | | | |
| Qy | 241 | GAACAGTTGCAGGGTCAAGAAACATGTGTACTTTGATGAGAGGTGTTTCAATCCACTTT | 300 | | | | | | | |
| Db | 279 | GAGCAGTTGCAGGGTCAAGAAACATGTGTACTTTATACGAGCTTTTACCATCACTTC | 338 | | | | | | | |
| Qy | 301 | TCTCCAAATGCGCATACGCTGTCTTGTCTTGAACTCCGTGATGTCAAGTGTTCGAAA | 360 | | | | | | | |
| Db | 339 | TCTCCAAATGCGCATACGCTGTCTTGTCTTGAACTCTCGATGTCAAGGTTTCGAAA | 398 | | | | | | | |
| Qy | 361 | TTCCTTATATATTCAGTCAACAGGGTTGTTCCATGTGATGATACAGATTGCGCCGA | 420 | | | | | | | |
| Db | 399 | TTCCTTATATATTCAGTCAACAGGGTTGTTCCATGTGATGATACAGACTGTCGAGA | 458 | | | | | | | |
| Qy | 421 | GATGTTTTGGAGGCCCTTGCTTTTCTTCATCATGAGGGGTATGTCCATGGGGAACCTCAA | 480 | | | | | | | |
| Db | 459 | GATGCTCTGGAGGCCCTTGCTTTTCTTCACCAAGAGGGGTATGTCCATGGGGAACCTCAA | 518 | | | | | | | |
| Qy | 481 | CCAAGTAACATATTTGTGAGTGCAGAGATGAATGTTTTAACTATGACTTTGACTT | 540 | | | | | | | |
| Db | 519 | CCAAGTAACATCTGTGAGTGGGAGAACGATGCTTTTAACTTATTTGACTTTGACTC | 578 | | | | | | | |
| Qy | 541 | AGCTTCAAAGAGGCATTCAGATGTATAATTTTACACAGCGGGTATCGGGCTTCCA | 600 | | | | | | | |
| Db | 579 | AGCTTCAAAGAGGCATTCAGAGTGAATATTTTACACAGCGGGTATTCGAGCTCT | 638 | | | | | | | |
| Qy | 601 | GAAGCAGAATTCGAAATTCCTTGGCCGAGGCTGGCTGCGAGAGTGTATACAGATGTACC | 660 | | | | | | | |
| Db | 639 | GAAGCAGAACTGCAAGATTCCTTGGCCGAGGCGGCTGCGAGTGTATACAGATGTACC | 698 | | | | | | | |
| Qy | 661 | TCAGCTGTGATCTGTGGAGCCTTAGAATCATTTTACTGAAATGTTCTCAGATGAAA | 720 | | | | | | | |
| Db | 699 | TCAGCTGTGATCTGTGGAGCCTTAGAATCATTTTACTGAAATGTTCTCAGATGAAA | 758 | | | | | | | |
| Qy | 721 | CTGAAATATCAGTCAGATCTCAGAGATTGAAGGCAACAGTCTGTATTTATTTAGTAC | 780 | | | | | | | |
| Db | 759 | CTGAAATATCAGTCAGATCTCAGAGATTGAAGGCAACAGTCTGTATTTATTTAGTAT | 818 | | | | | | | |
| Qy | 781 | ATATTTCCAGTAAAGCAGTGTGAATGCGGCAATTCAGCTTATCACTTAAGAGACTT | 840 | | | | | | | |
| Db | 819 | ATATTTCCAGTAAAGCAGTGTGAATGCGGCAATTCAGCTTATCACTTACAGAGACTT | 878 | | | | | | | |
| Qy | 841 | ATCAAAAGCATGCTTCAATGATCCAGCAGAAAGATTCCTGCTGAATGCGATGTGTC | 900 | | | | | | | |
| Db | 879 | ATCAAAAGCATGCTTCAATGACGCCAAGCAGAAAGATTCCTGCTGAATGCGCTTGTGC | 938 | | | | | | | |
| Qy | 901 | AGGCCATTTCTTATGATTCCTTTTGCCCGCATVTTGAAGTCTGTCATGCTCCACT | 960 | | | | | | | |
| Db | 939 | AGGCCATTTCTTATGATTCCTTTTGCCCGCATVTTGAAGTCTGTCATGCTCCACT | 998 | | | | | | | |
| Qy | 961 | CAAGTGTAAAGCTGTGAATGTGCTGAGATGATGATTTATCTTGGAAATGAAGAAATAT | 1020 | | | | | | | |

Db 999 CCAAGTCGACACTCCTCAATGCTGATGATGATCTTGTGAAATGAGATGATAT 1058
 QY 1021 GAAGATGTTAGAAAGATGTAAGAGAGAGTGTCAAAATATGACCGAGTATCTCTA 1080
 Db 1059 GAAGATGTTAGAAAGATGTAAGAGAGAGTGTCAAAATATGACCGAGTATCTCTG 1118
 QY 1081 CTGTGTTCCAAAGAAAATCTGTGACAGAGCAAGTCTTTGTTGATGTAAGAAATGCTGTG 1140
 Db 1119 CTGTGTTCCAAAGAAAATCTGTGACAGAGCAAGTCTTTGTTGATGTAAGAAATGCTGTG 1178
 QY 1141 GATTCACAAAGCTGCGCAAAATTAATCTGATGTAAGAGATGTTGATGTAAGAAATGCTGTG 1200
 Db 1179 GATTCACAAAGCTGCGCAAAATTAATCTGATGTAAGAGATGTTGATGTAAGAAATGCTGTG 1238
 QY 1201 GTACATTTACCGCGAGTGTGCTACAGAGAGGATATGATCAAACTTGTCTTAA 1260
 Db 1239 GCTACATTTACCGCGAGTGTGCTACAGAGAGGATATGATCAAACTTGTCTTAA 1298

RESULT 12 3244 bp mRNA linear ROD 08-SEP-1999
 Locus RNU70372
 DEFINITION Rattus norvegicus PAM COOH-terminal Interactor Protein 2 mRNA,
 complete cds.
 accession U70372.2 GI:5821767
 VERSION U70372
 KEYWORDS
 SOURCE
 ORGANISM Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 3244)
 Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
 and Bipper,B.A.
 Novel proteins that interact with the COOH-terminal cytosolic
 routing determinants of an integral membrane peptide-processing
 enzyme
 J. Biol. Chem. 271 (45), 28636-28640 (1996)

JOURNAL MEDLINE 97067094
 PUBMED 8910496
 2 (bases 1 to 3244)
 Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
 Bipper,B.A. and Mains,R.E.
 The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
 Determinants of the Peptide Processing Enzyme Peptidylglycine
 alpha-Aminidating Monooxygenase
 J. Biol. Chem. 274 (1999) In press
 3 (bases 1 to 3244)
 Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
 and Bipper,B.A.
 Direct Submission
 Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 4 (bases 1 to 3244)
 Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
 Bipper,B.A. and Mains,R.E.
 Direct Submission
 Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 REMARK
 COMMENT Sequence update by submitter
 FEATURES On Sep 3, 1999 this sequence version replaced gi:1698778.
 source location/Qualifiers
 1..3244
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 167..1421
 /function="serine protein kinase"
 /note="P-CIP2"
 /codon_start=1

/product="PAM COOH-terminal Interactor protein 2"
 /protein_id="AAC53031.2"
 /db_xref="GI:5821768"
 /translation="MAGSGCANGAPPRLEAFGLMVOYQSLSGSSASVYRACCG
 TPGSPGALKQPLPPTGTGAASABAYGSRRALEBDJQGRNIVLTVGFTIHSP
 NVPSRCLAELEADVSEILVYSSHQGSCSMNIQICARDVLEALFLHGVTHIDP
 PNIIWLSAENECEFKLIDFGLSPKEGQDVYVQTDGYRAPEALONCAQGLQSTPE
 CTSAVDLMELGIIILEMFSGMKLKHTVRSQEKANSASAIIDI PASKAVNNAIPAYH
 LADLIKSMHDDPSRRIRPAMALCSPPRSIIPAPHIEDLVMLPTVRLRLNLDYDL
 ENDEDEDVEDVEKRCQRYGQYVSLIPKKNPFGQYFVEYANAGDSAAQKLTLGR
 MPDGKRVATFPPLSAVKKRGLYQTLT"

ORIGIN
 Query Match 87.9%; Score 1108; DB 10; Length 3244;
 Best Local Similarity 92.5%; Pred. No. 26-276;
 Matches 1165; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 1 ATGGCGGGAATCCGGCTGCGCTGTGGGCGGAGCCGCGCTTTCTGTGAGGCTTCCGG 60
 Db 162 ATGGCGGGAATCCGGCTGCGCTGTGGGCGGAGCCGCGCTTCTGTGAGGCTTCCGG 221
 QY 61 CGGCTGTGGCAGGTACAGAGCCGCTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
 Db 222 CGGCTGTGGCAGGTACAGAGCCGCTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 281
 QY 121 CGCTGTGGGCAACCCCTGCGCCCGCGGCGGCTCAAGCAGTTCTTGCAGCAGGA 180
 Db 282 CGCTGTGGGCACTTCAGGCTGCGCCCGCGGCGGCTCAAGCAGTTCTTGCAGCAGGA 341
 QY 181 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGCGCTG 240
 Db 342 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGCGCTG 401
 QY 241 GAACAGTTGCAAGGTACAGAAACATCGTACTTTGATGAGTGTATCAATCACTTT 300
 Db 402 GAGCAGTTGCAAGGTACAGAAACATCGTACTTTGATGAGTGTATCAATCACTTT 461
 QY 301 TCTCCAAATGTCATCAGCTGTCTGTGCTTGAACCTCGTAGATGTCAGTCTTCGGA 360
 Db 462 TCTCCAAATGTCATCAGCTGTCTGTGCTTGAACCTCGTAGATGTCAGTCTTCGGA 521
 QY 361 TTGCTCTTATTTCCAGTACACAGGGTGTGTTCCATGTGATGATACAGATTCGCCGA 420
 Db 522 TTGCTCTTATTTCCAGTACACAGGGTGTGTTCCATGTGATGATACAGATTCGCCGA 581
 QY 421 GATGTTTGAAGCCCTTCTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 480
 Db 582 GATGTTTGAAGCCCTTCTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 641
 QY 481 CCAAGTAATATGATGAGTGCAGAGATGAATGTTTAACTCATGACTTGAAGCTT 540
 Db 642 CCAAGTAATATGATGAGTGCAGAGATGAATGTTTAACTCATGACTTGAAGCTT 701
 QY 541 ACCTTCAAGAGCAATCAGATGTAAAGTATTTACAGACAGAGGATTCGGCTCA 600
 Db 702 ACCTTCAAGAGCAATCAGATGTAAAGTATTTACAGACAGAGGATTCGGCTCA 761
 QY 601 GAAGCAGAAATTCGAAATGCTTGGCCCGAGGCTGCGCTGAGATGATACAGATGTACC 660
 Db 762 GAAGCAGAAATTCGAAATGCTTGGCCCGAGGCTGCGCTGAGATGATACAGATGTACC 821
 QY 661 TCAGCTGTGATCTGTGAGCCTTGAAGATCATTTTACAGAAATGTTCTCAGAAATGAAA 720
 Db 822 TCAGCTGTGATCTGTGAGCCTTGAAGATCATTTTACAGAAATGTTCTCAGAAATGAAA 881
 QY 721 CTGAAACATACAGTACATCTCAGAAATGAGAGCAACAGTTCTGTATTTATGATCAC 780
 Db 882 CTGAAACATACAGTACATCTCAGAAATGAGAGCAACAGTTCTGTATTTATGATCAT 941
 QY 781 AATATTGCGAATGAAAGCAGTGTGATGCGCAATTCAGCTTACCTTAAGAGACTT 840
 Db 942 AATATTGCGAATGAAAGCAGTGTGATGCGCAATTCAGCTTACCTTAAGAGACTT 1001

| | | | | |
|------------------|--|------|--|-----------------------------|
| QY | | 841 | ATCAAAAGCATGCTTCATGAATGCATCCAAAGCAGAAGAAATTCCGGCTGAATGGCAATTGTGC | 900 |
| Dd | | 1002 | ATCAAAGACATGCTTCAATGACGACCAGCAAGAGAACCTCGCTGAGATGGCCTTGTC | 1061 |
| QY | | 901 | AGCCATCTCTTAGCATTCCTTTTGGCCCCCTCATATTGAAGATCTGTGATGCTCCACT | 960 |
| Dd | | 1062 | AGGCCATTCTTAGATTCCTTTTGGCCCTCATATTGAAGATCTGTGATGCTCCAACT | 1121 |
| QY | | 961 | CCAAGTCTAGACCTGCTGATATGTGCTGGATGATGATTATCTTTGGGAATGAAGAGGAATAT | 1020 |
| Dd | | 1122 | CCAGTGTCTACAACCTCCCAATGTGCTGATGATGATCTATCTTGAATAAAGAGATGAATAT | 1181 |
| QY | | 1021 | GAAATGTTGTAGAAGATGTAAAAGAGAGTGTCAAATAATATGACCGAGTGTATCTCTA | 1080 |
| Dd | | 1182 | GAAATGTTGTAGAAGATGTAAAAGAGAGTGTCAAATAATATGACCGAGTGTATCTCTG | 1241 |
| QY | | 1081 | CTTGTTCCAAAGAAATCTGTGCAGAGACAAAGTCTTTGTGTGATATGCAATGCTGGT | 1140 |
| Dd | | 1242 | CTTGTTCCAAAGAAATCTGTGCAGAGACAAAGTCTTTGTGTGATATGCAATGCTGGT | 1301 |
| QY | | 1141 | GATTCCAAAGCTGCGCAGAAATTACTGACTGGAMGATGTTTGAATGGGAAGTTGTGTG | 1200 |
| Dd | | 1302 | GATTCCAAAGCTGCGCTCAGAAATGCTGACTGGAGAGATGTTTGAACGGAAGTTGTGTG | 1361 |
| QY | | 1201 | GCTACATTCTACCCGCTGAGTGCCTPACAAGAGGGGATATCTGTATCAAACCTGCTTAA | 1260 |
| Dd | | 1362 | GCTACATTCTACCCGCTGAGTGCCTPACAAGAGGGGATATCTTATCAAACCTGCTTAA | 1421 |
| RESULT_13 | | | | |
| MMPKIS | | | | |
| LOCUS | | | 1254 bp | mRNA linear ROD 08-JUL-2002 |
| DEFINITION | | | Mus musculus partial mRNA for Kis protein kinase (Kis gene). | |
| ACCESSION | | | X10725 | |
| VERSION | | | X10725.2 | GI:21726712 |
| KEYWORDS | | | Kis gene; protein kinase. | |
| SOURCE | | | Mus musculus (house mouse) | |
| ORGANISM | | | Mus musculus | |
| | | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| | | | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | |
| | | | 1 (bases 1 to 1254) | |
| REFERENCE | | | Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and | |
| AUTHORS | | | Sobel,A. | |
| TITLE | | | Kis is a protein kinase with an RNA recognition motif | |
| JOURNAL | | | J. Biol. Chem. 272 (37), 23151-23156 (1997) | |
| MEDLINE | | | 97435279 | |
| PUBMED | | | 9287318 | |
| REFERENCE | | | 2 | |
| AUTHORS | | | Maucuer,A. | |
| TITLE | | | Direct Submission | |
| JOURNAL | | | Submitted (23-JUN-1997) Maucuer A., INSERM U440, U440, 17 Rue du | |
| REMARK | | | Fer Moulin, Paris, 75005, FRANCE | |
| AUTHORS | | | revised by [4] | |
| TITLE | | | 3 (bases 1 to 1254) | |
| JOURNAL | | | Maucuer,A. | |
| REMARK | | | Direct Submission | |
| AUTHORS | | | Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du | |
| TITLE | | | Fer Moulin, Paris, 75005, FRANCE | |
| JOURNAL | | | On Jul 10, 2002 this sequence version replaced gi:1806129. | |
| COMMENT | | | Overlaps with related sequences X82320, X10725. | |
| FEATURES | | | location/Qualifiers | |
| SOURCE | | | 1..1254 | |
| | | | /organism="Mus musculus" | |
| | | | /mol_type="mRNA" | |
| | | | /db_xref="taxon:10090" | |
| | | | /clone="Xk23" | |
| | | | /tissue_type="brain" | |
| | | | /clone_id="lambda gt10" | |
| | | | /dev_stage="neonate" | |
| | | | 1..1254 | |
| | | | /gene="Kis" | |
| | | | 11..>1254 | |
| | | | /gene="Kis" | |
| gene | | | | |
| CDS | | | | |

[illegible]


```

Db      791 ATATTGCGAGTAACAGAGTGGATGCGGCAATTCAGCCATACCTCAGAGACCTT 850
Qy      841 ATCAAAAGCATGCTTATGATGATCCAGAGAAATTCCTGTAATGCGATTGTC 900
Db      851 ATCAAAAGCATGCTTATGATGATCCGCGGAGAAATCCCTGATGATGCGATTGTC 910
Qy      901 AGCCCATCTTATGATGATGCTTGGCCCTCATATGAAAGATGCTGATGCTCCACT 960
Db      911 AGCCCATCTTATGATGATGCTTGGCCCTCATATGAAAGATGCTGATGCTCCACT 970
Qy      961 CCAGTCTAAGCATGCTGATGATGCTGATGATGATGATGATGATGATGATGAT 1020
Db      971 CCAGTCTAAGCATGCTGATGATGCTGATGATGATGATGATGATGATGATGAT 1030
Qy      1021 GAAGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db      1031 GAAGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1090
Qy      1081 CTGTTCCAAAGGAAATCTGCGAGAGACAAAGTCTTGTGATGATGATGATGATGAT 1140
Db      1091 CTGTTCCAAAGGAAATCTGCGAGAGACAAAGTCTTGTGATGATGATGATGATGAT 1150
Qy      1141 GATTCGAAGCTGCGAGAAATTAATGATGATGATGATGATGATGATGATGATGATG 1200
Db      1151 GATTCGAAGCTGCGAGAAATTAATGATGATGATGATGATGATGATGATGATGATG 1210
Qy      1201 GCTACATTTACCCGCTGAGTGCCTCAAGAGGGATATCTGTA 1244
Db      1211 GCTACATTTACCCGCTGAGTGCCTCAAGAGGGATATCTGTA 1254

```

RESULT 14

```

LOCUS      BC026046
DEFINITION Homo sapiens kinase interacting with leukemia-associated gene
            (src1hmin), mRNA (CDNA clone MGC:39391 IMAGE:4649261), complete
            cds.
ACCESSION  BC026046
VERSION    BC026046.1 GI:19684094
KEYWORDS   MGC.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE 1 (bases 1 to 1367)
Strausberg, R.L., Feingold, D.A., Grouse, L.H., Derge, J.G., Klusner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stabileton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedlin, T.B., Toshikiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McKean, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, X., Gibbs, R.A., Viallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E., Butcher, A., Schein, J.E., Jones, S.J., Jones, S.J., Smallus, D.E.,

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL PUMED 12477932
REFERENCE 2 (bases 1 to 1367)
TITLE Strausberg, R.
AUTHORS Direct Submission
JOURNAL Submitted (20-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer

REMARK

COMMENT

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chow, William Crow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Maason, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabh, Parvaneh Saeedi, JR Santos, Angelique Scherch, Ursula Skalska, Duane Smalins, Jeff Stott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.

FEATURES

source

1..1367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:39391 IMAGE:4649261"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..1367
/gene="KIS"
/db_xref="LocusID:127933"
159..1193
/gene="KIS"
/codon_start=1
/product="kinase interacting src1hmin"
/protein_id="AAH26046.1"
/db_xref="GI:19684095"
/db_xref="LocusID:127933"
/translation="MAGSGCAGAPPRFLAFLGLWQVSRSGSSASVYRVCCG
NPGSPGALKQFLPGTTGAASAEYFRERAAIBOLQGRNIVTLVGVFTIFSP
NVPSCLELIDVSELILYSSHOGSMWIOHCARVLEALFLHREGVHADLK
PNTLSAENECFKIDRGLSPKRNODVKTIOGDVAPAEALONCLIAOGLSDTE
CTSAVDMSLIGTILHMFSGMKLTATVSGEKANSATIDHIFPSKAVVNAIATYH
LNDLTKSLHDDPSRRTFAEVALCSPPFSIPAPHTEDLVMLPTVLRNLNLDYLL
ENEREYELIC"

CDS

gene

ORIGIN

Query Match 84.7%; Score 1067.2; DB 9; Length 1367;
Best Local Similarity 92.7%; Pred. No. 6,5e-266;
Matches 1168; Conservative 0; Mismatches 3; Indels 89; Gaps 1;
Qy 1 ATGGCGGATCCGCGCTGCGCTGCGCGCGAGCGCGCGCTTTCTGAGAGCCTTCGG 60
Db 159 ATGGCGGATCCGCGCTGCGCTGCGCGCGAGCGCGCGCTTTCTGAGAGCCTTCGG 218
Qy 61 CGGCTGTGGCAGTACAGAGCCGCTGCGGTAGCGGCTCTCGGCTCGGTGTAGGGTT 120
Db 219 CGGCTGTGGCAGTACAGAGCCGCTGCGGTAGCGGCTCTCGGCTCGGTGTAGGGTT 278
Qy 121 CGGCTGTGGCAGTACAGAGCCGCGCGCGCGCGCGCTCAAGCAGTCTTGGCGGCGAG 180
Db 279 CGGCTGTGGCAGTACAGAGCCGCGCGCGCGCGCGCTCAAGCAGTCTTGGCGGCGAG 338
Qy 181 ACCACCGGGGCTGCGCGCTTGGCGCGCGAGATAGTTTCCGAAAGAGAGGGCGGCTG 240
Db 339 ACCACCGGGGCTGCGCGCTTGGCGCGCGAGATAGTTTCCGAAAGAGAGGGCGGCTG 398


```
Db 423 AAGTATATTCAAGACAGCGGTATCGGGCTCCAGAACAGAAATTGC AAAATTGCTTGCC 482
Qy 628 CAGGCTGGCTGTGAGAGTATACAGAAATGACTCAGCTGTGATCTGTGGAGCCTAGGA 687
Db 483 CAGGCTGGCTGTGAGAGTATACAGAAATGACTCAGCTGTGATCTGTGGAGCCTAGGA 542
Qy 688 ATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAATCTCAGGAA 747
Db 543 ATCATTTTACTGGAAATGTTCTCAGGAATGAAACTGAAACATACAGTCAATCTCAGGAA 602
Qy 748 TGGAGGCAAAACAGTTCTGCTATTATTGATCAGATATTGCCAGTAAGCAGTGTGAAT 807
Db 603 TGGAGGCAAAACAGTTCTGCTATTATTGATCAGATATTGCCAGTAAGCAGTGTGAAT 662
Qy 808 GCCGCAATTCAGCCCTATCACCCTTAAGAGCCTTATCAAAAAGCATGCTTCATGATGCCA 867
Db 663 GCCGCAATTCAGCCCTATCACCCTTAAGAGCCTTATCAAAAAGCATGATGATGCCA 722
Qy 868 AGCAGAAAGATTCTGCTGTAATGCGATTGTGACGCCCATCTTTAGCATTCCTTTGCG 927
Db 723 AGCAGAAAGATTCTGCTGTAATGCGATTGTGACGCCCATCTTTAGCATTCCTTTGCG 782
Qy 928 CCTCATATTGAAGATCTGTCATGCTTCCACTCCAGTGTMAAGCTGCTGAATGTGCTG 987
Db 783 CCTCATATTGAAGATCTGTCATGCTTCCACTCCAGTGTMAAGCTGCTGAATGTGCTG 842
Qy 988 GATGATGATTATCTTGGGAATGAAGAGAAATGAAGATGTTGTAGAAGATGTAAAGAG 1047
Db 843 GATGATGATTATCTTGGGAATGAAGAGAAATGAAGATGTTGTAGAAGATGTAAAGAG 902
Qy 1048 GAGTGTCAAAAATATGACCAAGTGTATCTCTACTTGTCCAAAAGAAAATCCTGSCAGA 1107
Db 903 GAGTGTCAAAAATATGACCAAGTGTATCTCTACTTGTCCAAAAGAAAATCCTGSCAGA 962
Qy 1108 GGAACAAGTCTTGTGATGCAAAATGCTGTGATTCCAAAGCTGCGCAAAAATTACTG 1167
Db 963 GGAACAAGTCTTGTGATGCAAAATGCTGTGATTCCAAAGCTGCGCAAAAATTACTG 1022
Qy 1168 ACTGGAAGATGTTNGATGGGAAGTTGTTGGGCTACATTCACCCGCTGAGTGCCTAC 1227
Db 1023 ACTGGAAGATGTTNGATGGGAAGTTGTTGGGCTACATTCACCCGCTGAGTGCCTAC 1082
Qy 1228 AAGAGGGATATCTGTATCAAACTTGCTTTAA 1260
Db 1083 AAGAGGGATATCTGTATCAAACTTGCTTTAA 1115
```

Search completed: November 29, 2004, 14:49:16
Job time : 5517 secs

This Page Blank (uspto)

XX The present DNA sequence encodes a wild type human KIS (hKIS), isolated
 CC from a human B-cell library, using a yeast two hybrid screening system.
 CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
 CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
 CC that controls cell proliferation and is localised predominantly in the
 CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
 CC G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
 CC The hKIS sequences are used to modulate cell proliferation and treat cell
 CC proliferative and vascular diseases. The polynucleotide sequence may be
 CC used in gene therapy to treat vascular disorders such as restenosis or
 CC atherosclerosis
 XX

SQ Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

Query Match 100.0%; Score 1260; DB 3; Length 1260;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGATCCGCGTGCCTGGGGGCGGAGCCGCGCGTTTCTGAGAGCCTTCGGG 60
 Db 1 ATGGCGGATCCGCGTGCCTGGGGGCGGAGCCGCGCGTTTCTGAGAGCCTTCGGG 60
 QY 61 CGGCTGTGGGAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
 Db 61 CGGCTGTGGGAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
 QY 121 CGCTGTGGCGGCAACCCCTGGCTGCCCCCGGCGCCCTCAAGAGATTCTTCCGCCAGGA 180
 Db 121 CGCTGTGGCGGCAACCCCTGGCTGCCCCCGGCGCCCTCAAGAGATTCTTCCGCCAGGA 180
 QY 181 ACCAGCGGGGCTGCGGCTCTGCGCGGAGTATGTTTCCGAAAGAGGGCGGCTG 240
 Db 181 ACCAGCGGGGCTGCGGCTCTGCGCGGAGTATGTTTCCGAAAGAGGGCGGCTG 240
 QY 241 GAACATTTGCAAGGTACAGAGAACATCGTACTTTGTATGAGATGTTTCAATCCACTTT 300
 Db 241 GAACATTTGCAAGGTACAGAGAACATCGTACTTTGTATGAGATGTTTCAATCCACTTT 300
 QY 301 TCTCCAAATGTGCGATCAGCTGTCTGTTGCTTGAATCCCTGAGATCAGTGTTCGGAA 360
 Db 301 TCTCCAAATGTGCGATCAGCTGTCTGTTGCTTGAATCCCTGAGATCAGTGTTCGGAA 360
 QY 361 TTGCTCTTATATTCAGATCAGAGGGTGTCTTCATGTGATGATACAGATTCGCCCGA 420
 Db 361 TTGCTCTTATATTCAGATCAGAGGGTGTCTTCATGTGATGATACAGATTCGCCCGA 420
 QY 421 GATGTTTGGAGGCCCTTGCTTTTCTCATGAGAGGCTATGTCATGCGGACCTCAAA 480
 Db 421 GATGTTTGGAGGCCCTTGCTTTTCTCATGAGAGGCTATGTCATGCGGACCTCAAA 480
 QY 481 CCACGTATATTTGAGAGTGCAGAGAAATGATGTTTAACTCAATTGACTTGGACTT 540
 Db 481 CCACGTATATTTGAGAGTGCAGAGAAATGATGTTTAACTCAATTGACTTGGACTT 540
 QY 541 AGCTTCAAAAGAGGCAATCAGATGTAAGTATATTCAGACAGACGGGTATCGGGCTCA 600
 Db 541 AGCTTCAAAAGAGGCAATCAGATGTAAGTATATTCAGACAGACGGGTATCGGGCTCA 600
 QY 601 GAAGCAGATTTGCAAAATGCTTGGCCGAGGCTGCGCTGAGAGTATACAGAAATGATCC 660
 Db 601 GAAGCAGATTTGCAAAATGCTTGGCCGAGGCTGCGCTGAGAGTATACAGAAATGATCC 660
 QY 661 TTAGCTGTGATCTGTGAGAGCTAGAGATCATTTTACTGGAATGTTTCCAGGAATGAAA 720
 Db 661 TTAGCTGTGATCTGTGAGAGCTAGAGATCATTTTACTGGAATGTTTCCAGGAATGAAA 720
 QY 721 CTGAAACATACACTCAGATCTCAGAAATGAGGCAAAAGTTCTCTATATATGATCAC 780
 Db 721 CTGAAACATACACTCAGATCTCAGAAATGAGGCAAAAGTTCTCTATATATGATCAC 780
 QY 781 AATATTTCCAGTAAACAGATGTGATGCGCAATTCAGCCATACCTAAGAGCCTT 840
 Db 781 AATATTTCCAGTAAACAGATGTGATGCGCAATTCAGCCATACCTAAGAGCCTT 840

Db 781 AATATTTCCAGTAAACAGATGTGATGCGCAATTCAGCCATACCTAAGAGCCTT 840
 QY 841 ATCAAAAGCATCTTCATGATGATCCAGCAGAGAAATCTGCTGAAATGSCATTGTC 900
 Db 841 ATCAAAAGCATCTTCATGATGATCCAGCAGAGAAATCTGCTGAAATGSCATTGTC 900
 QY 901 AGCCCATTTTTCAGATTCCTTTTGCCCTCATATTTGAAGATCTGTCATGCTCCACT 960
 Db 901 AGCCCATTTTTCAGATTCCTTTTGCCCTCATATTTGAAGATCTGTCATGCTCCACT 960
 QY 961 CCAGTCTTAAAGCTGTAATGTGCTGATGATGATTTATCTTGGAAATCAAGAGAAAT 1020
 Db 961 CCAGTCTTAAAGCTGTAATGTGCTGATGATGATTTATCTTGGAAATCAAGAGAAAT 1020
 QY 1021 GAAGATGTTGTAAGATGTAAGAGAGAGTCTCAAAATATGAGCAAGTGTATCTCTA 1080
 Db 1021 GAAGATGTTGTAAGATGTAAGAGAGAGTCTCAAAATATGAGCAAGTGTATCTCTA 1080
 QY 1081 CTTGTTCCAAAGAAATCTGCGAGAGGACAAATCTTTGTTGAGTATGCAAAATCTGGT 1140
 Db 1081 CTTGTTCCAAAGAAATCTGCGAGAGGACAAATCTTTGTTGAGTATGCAAAATCTGGT 1140
 QY 1141 GATTCCAAGCTGCGCAGAAATTAAGTCTGAGAGAGTGTGATGAGGAAATGTTGTG 1200
 Db 1141 GATTCCAAGCTGCGCAGAAATTAAGTCTGAGAGAGTGTGATGAGGAAATGTTGTG 1200
 QY 1201 GCTACATTTACCCCTGAGTCTCTACAGAGGGATATCTGTATCAAACTTGTCTTAA 1260
 Db 1201 GCTACATTTACCCCTGAGTCTCTACAGAGGGATATCTGTATCAAACTTGTCTTAA 1260

RESULT 2
 AA251356
 ID AA251356 standard; DNA; 1260 BP.
 XX
 AC AA251356;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.
 XX
 KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
 KW transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27;
 KW modulator; treatment; cell proliferative disease; vascular disorder;
 KW gene therapy; atherosclerosis; restenosis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..1260
 FT CDS
 FT /*tag= a
 FT /product= "Human transdominant mutant KIS protein"
 FT /transl_except= (pos:556..561, Asn)
 FT misc_feature
 FT 161
 FT /*tag= b
 FT /note= "A to G substitution resulting in a Lys to Arg
 change in the amino acid sequence"
 FT
 FT
 PN MO20001165-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 20-AUG-1999; 99WO-US018903.
 XX
 PR 21-AUG-1998; 98US-0097710P.
 XX
 PA (NABE/) NABEL G J.
 PA (NABE/) NABEL E G.
 XX
 PI Nabel GJ, Nabel EG;
 XX
 DR MPI: 2000-237648/20.
 DR P-PSDB; AAT70306.

XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
 PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 PT proliferation.

XX Claim 7, Page 59; 70pp; English.

XX The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
 CC constructed by site directed mutagenesis. A single nucleotide
 CC substitution (A to G) results in a lysine to arginine change in the
 CC protein sequence. hKIS is a serine/threonine kinase, that acts as an
 CC inhibitor of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
 CC controls cell proliferation and is localised predominantly in the
 CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
 CC G1 phase. The hKIS sequences are used to modulate cell proliferation and
 CC treat cell proliferation and vascular diseases. The polynucleotide
 CC sequence may be used in gene therapy to treat vascular disorders such as
 CC restenosis or atherosclerosis

XX Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

XX Query Match 99.9%; Score 1258.4; DB 3; Length 1260;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 1259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGGATCCGGCTGGCGGCGGCGGAGCGCGCGCTTTTGGAGGCGCTTCGGG 60
 DB 1 ATGGCGGGATCCGGCTGGCGGCGGCGGAGCGCGCGCTTTTGGAGGCGCTTCGGG 60
 QY 61 CGGCTGTGGCAGGTACAGAGCCGTGTGGGTAGCGGCTCCGCTCGGCTGTATCGGGTT 120
 DB 61 CGGCTGTGGCAGGTACAGAGCCGTGTGGGTAGCGGCTCCGCTCGGCTGTATCGGGTT 120
 QY 121 CGCTGTGCGGCAACCTTGCTGCGCCCGCGGCGCTTCAAGAGTTCTTGGCCGAGGA 180
 DB 121 CGCTGTGCGGCAACCTTGCTGCGCCCGCGGCGCTTCAAGAGTTCTTGGCCGAGGA 180
 QY 181 ACCACCGGGGCTGCGGCGCTTGCAGGATGTTTCCGAAAGAGAGGGGCGGCTG 240
 DB 181 ACCACCGGGGCTGCGGCGCTTGCAGGATGTTTCCGAAAGAGAGGGGCGGCTG 240
 QY 241 GAACAGTTGACAGGTACAGAAACATCGTGAATTTGTATGAGTGTTCATCACTTT 300
 DB 241 GAACAGTTGACAGGTACAGAAACATCGTGAATTTGTATGAGTGTTCATCACTTT 300
 QY 301 TCTCCAAATGTGCATCAGCTGTCTGTGCTTGAATCTCTGAGTCAAGTGTTCGAA 360
 DB 301 TCTCCAAATGTGCATCAGCTGTCTGTGCTTGAATCTCTGAGTCAAGTGTTCGAA 360
 QY 361 TTGCTCTTATTTCCAGTACACGAGGTTTTCATGTGATGATTAAGCATTTGGCCGCA 420
 DB 361 TTGCTCTTATTTCCAGTACACGAGGTTTTCATGTGATGATTAAGCATTTGGCCGCA 420
 QY 421 GATGTTTGGAGGCGCTTCTTCTTCATCATGAGGCTATGCTCATGCGACCTCAAA 480
 DB 421 GATGTTTGGAGGCGCTTCTTCTTCATCATGAGGCTATGCTCATGCGACCTCAAA 480
 QY 481 CCAAGTAATTTGGAGTGCAGAGATGAATGTTTAACTCAATTGACTTGAATT 540
 DB 481 CCAAGTAATTTGGAGTGCAGAGATGAATGTTTAACTCAATTGACTTGAATT 540
 QY 541 AGCTTAAAGAGAGCATCAGATGTAAGTATATTCAGACAGGATTCGGGCTCCA 600
 DB 541 AGCTTAAAGAGAGCATCAGATGTAAGTATATTCAGACAGGATTCGGGCTCCA 600
 QY 601 GAAGCAGAAATTCGAAATTCCTTGCCAGGCTGCGCTGAGAGTATACAGATATACC 660
 DB 601 GAAGCAGAAATTCGAAATTCCTTGCCAGGCTGCGCTGAGAGTATACAGATATACC 660
 QY 661 TCAGCTGTGATGTGGAGCTTGAAGATCAATTTTACTGGAATGTTTTCAGAAATGAAA 720
 DB 661 TCAGCTGTGATGTGGAGCTTGAAGATCAATTTTACTGGAATGTTTTCAGAAATGAAA 720

QY 721 CTGAACATACAGTCAGATCTCAGGAATGAGCAAGCAAGCTTCTGATTTATTTGATC 780
 DB 721 CTGAACATACAGTCAGATCTCAGGAATGAGCAAGCAAGCTTCTGATTTATTTGATC 780
 QY 781 ATATTGGCAGTAAAGCAGTGTGAATCCGCAATTCAGCCCTATCAGCTTAAGACCTT 840
 DB 781 ATATTGGCAGTAAAGCAGTGTGAATCCGCAATTCAGCCCTATCAGCTTAAGACCTT 840
 QY 841 ATCAAAAGCATGCTTCATGATGATTCAGAGCAAGAAATTCCTGCTGGAATGCGATTGCG 900
 DB 841 ATCAAAAGCATGCTTCATGATGATTCAGAGCAAGAAATTCCTGCTGGAATGCGATTGCG 900
 QY 901 AGCCATCTTTAGGATTCCTTTTCCCTCATATTTGAAGATCTGCTATGCTTCCACT 960
 DB 901 AGCCATCTTTAGGATTCCTTTTCCCTCATATTTGAAGATCTGCTATGCTTCCACT 960
 QY 961 CCAAGTCTAAGCATGCTGAATGCTGTGATGATTTATCTTTGGAGATGAAGAGAAAT 1020
 DB 961 CCAAGTCTAAGCATGCTGAATGCTGTGATGATTTATCTTTGGAGATGAAGAGAAAT 1020
 QY 1021 GAAGATGTTTGAAGATGTAAGAGAGATGTCAAAATATGACCAAGTGTATCTCTA 1080
 DB 1021 GAAGATGTTTGAAGATGTAAGAGAGATGTCAAAATATGACCAAGTGTATCTCTA 1080
 QY 1081 CTTGTTCCAAAGAAATCTCTGCGAGAGCAAGCTTTGTTGATGATGCAAAATCTGGT 1140
 DB 1081 CTTGTTCCAAAGAAATCTCTGCGAGAGCAAGCTTTGTTGATGATGCAAAATCTGGT 1140
 QY 1141 GATTCCAAAGCTGCGCAGAAATTAAGTCTGAGAGATGTTTGAATGAGAAATTTGTTG 1200
 DB 1141 GATTCCAAAGCTGCGCAGAAATTAAGTCTGAGAGATGTTTGAATGAGAAATTTGTTG 1200
 QY 1201 GCTACATTTCTACCCGCTGAGGCTTCAAGAGGGGATATCTGATCAAACTTGTCTTAA 1260
 DB 1201 GCTACATTTCTACCCGCTGAGGCTTCAAGAGGGGATATCTGATCAAACTTGTCTTAA 1260

RESULT 3

AB099383
 ID AB099383 standard; cDNA; 2008 BP.

XX AB099383;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 XX

DE Human coding sequence SEQ ID 116.

XX Human; expressed sequence tag; EST; haematopoietic disorder;
 KW central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; neutrotropic;
 KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective; gene therapy; gene; ss.

OS Homo sapiens.

PN WO200259260-A2.

PD 01-AUG-2002.

PF 16-NOV-2001; 2001WO-US042950.

PR 17-NOV-2000; 2000US-00714936.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;

PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI; 2002-590824/63.
 DR N-PSDB; ABP64797.

XX New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.

PS Claim 1; SEQ ID NO 116; 394bp; English.

XX The present invention relates to novel human coding sequences (AB099268-
 CC AB099608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotent or pluripotent state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcr_sequences

SO Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

Query Match 99.9%; Score 1258.4; DB 6; Length 2008;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGGAGATCCGGCTGCGGCTGGGGCGGGAGACCGCGGCTTTTGAGAGCGCTTCGGG 60
 DB 79 ATGGCGGAGATCCGGCTGCGGCTGGGGCGGGAGACCGCGGCTTTTGAGAGCGCTTCGGG 138
 QY 61 CGGCTGTGGAGGTACAGAGCCGCTCTGGTGGAGCGGCTCTCGCTCGGTATTCGGGATT 120
 DB 139 CGGCTGTGGAGGTACAGAGCCGCTCTGGTGGAGCGGCTCTCGCTCGGTATTCGGGATT 198
 QY 121 CGGCTGTGGAGGTACAGAGCCGCTCTGGTGGAGCGGCTCTCGCTCGGTATTCGGGATT 180
 DB 199 CGGCTGTGGAGGTACAGAGCCGCTCTGGTGGAGCGGCTCTCGCTCGGTATTCGGGATT 258
 QY 181 ACCACCGGGGCTGCGGCTCTGCGGCGAGATGTTTCCGCAAGAGAGGGCGGCGCTG 240
 DB 259 ACCACCGGGGCTGCGGCTCTGCGGCGAGATGTTTCCGCAAGAGAGGGCGGCGCTG 318
 QY 241 GAACAGTTGAGGGTACAGAAACATCGTACTTTGTATGAGAGTGTTCATCACTCTT 300
 DB 319 GAACAGTTGAGGGTACAGAAACATCGTACTTTGTATGAGAGTGTTCATCACTCTT 378
 QY 301 TCTCCAAATGTGTCATCAGCGCTGTCTGTGGTGAACCTCTGATGTCACTGTTTGGAA 360
 DB 379 TCTCCAAATGTGTCATCAGCGCTGTCTGTGGTGAACCTCTGATGTCACTGTTTGGAA 438
 QY 361 TTGCTCTTATTCAGAGTACAGAGGTTGTTCATGAGATGATAGAGATTGGCGCGA 420
 DB 439 TTGCTCTTATTCAGAGTACAGAGGTTGTTCATGAGATGATAGAGATTGGCGCGA 498
 QY 421 GATGTTTGGAGCCCTTGCTTTCTTCATCATGAGGCGTATGTCATGCGAGCTCAAA 480
 DB 499 GATGTTTGGAGCCCTTGCTTTCTTCATCATGAGGCGTATGTCATGCGAGCTCAAA 558
 QY 481 CCAAGTAACATATTGAGAGTGCAGAGAAATGATTTTAACTCATTTGAGACTT 540

DB 559 CCACCTAACATATTGTGGAGTGTCAGAGAAATGATGATTTTAACTCATTTGAGACTT 618
 QY 541 AGCTTCAAGAGGGAATCAGAGATGTAAGTATTCAGACAGAGGGATTCGGCTCA 600
 DB 619 AGCTTCAAGAGGGAATCAGAGATGTAAGTATTCAGACAGAGGGATTCGGCTCA 678
 QY 601 GAAGCAGATTTGCAAAATTTGCTTGCGCCAGGCTGCGCAGAGTGTATACAGATGTAAC 660
 DB 679 GAAGCAGATTTGCAAAATTTGCTTGCGCCAGGCTGCGCAGAGTGTATACAGATGTAAC 738
 QY 661 TAGCTGTGATCTGTGAGACCTTAGGAATCATTTTACTGGAATGTTTCTCAGGAATGAA 720
 DB 739 TAGCTGTGATCTGTGAGACCTTAGGAATCATTTTACTGGAATGTTTCTCAGGAATGAA 798
 QY 721 CTGAACATACAGTACAGATCTCAGGAATGGAAGCAACAGTTCTGATTTATGATCAG 780
 DB 799 CTGAACATACAGTACAGATCTCAGGAATGGAAGCAACAGTTCTGATTTATGATCAG 858
 QY 781 ATATTGCGCAGTAAAGCAGTGTGTAATGCGCAATTCAGAGCTTATCAGTAAAGACCTT 840
 DB 859 ATATTGCGCAGTAAAGCAGTGTGTAATGCGCAATTCAGAGCTTATCAGTAAAGACCTT 918
 QY 841 ATCAAAAGATGCTTCTGATGATATCCAGCAGAGAAATTCCTGCTGAATGSCATTGTC 900
 DB 919 ATCAAAAGATGCTTCTGATGATATCCAGCAGAGAAATTCCTGCTGAATGSCATTGTC 978
 QY 901 ACCCATTTCTTAGCATTCCTTTGCCCTCATATTGGAAGATCTGTCATGCTTCCACT 960
 DB 979 ACCCATTTCTTAGCATTCCTTTGCCCTCATATTGGAAGATCTGTCATGCTTCCACT 1038
 QY 961 CCAAGTCTAAGACTCTGATGTCGTGATGATATATCTTGGGAATGAAGAGAAATAT 1020
 DB 1039 CCAAGTCTAAGACTCTGATGTCGTGATGATATATCTTGGGAATGAAGAGAAATAT 1098
 QY 1021 GAAAGTGTGTGAAGATGTAAGAGAGTCTCAAAATATGACAGCTGATCTCTA 1080
 DB 1099 GAAAGTGTGTGAAGATGTAAGAGAGTCTCAAAATATGACAGCTGATCTCTA 1158
 QY 1081 CTTGTTCCAAGGAATCTCTGCGAGAGACAAAGTCTTGTGTGATATGCAAAATGCTGT 1140
 DB 1159 CTTGTTCCAAGGAATCTCTGCGAGAGACAAAGTCTTGTGTGATATGCAAAATGCTGT 1218
 QY 1141 GATTCGAAAGCTGCGCAGAAATTAAGTCTGAGAGAGTGTGATGGAAGTTGTTGTG 1200
 DB 1219 GATTCGAAAGCTGCGCAGAAATTAAGTCTGAGAGAGTGTGATGGAAGTTGTTGTG 1278
 QY 1201 GCTACATTTCTACCGCTGAGGCTTACCAAGAGGGGATATCTGTATCAAACTTGCTTAA 1260
 DB 1279 GCTACATTTCTACCGCTGAGGCTTACCAAGAGGGGATATCTGTATCAAACTTGCTTAA 1338

RESULT 4
 AA166829
 ID AA166829 standard; cDNA; 1260 BP.
 XX
 AC AA166829;
 XX
 DT 07-JAN-2002 (first entry)
 XX
 DE Human protein kinase polypeptide 13887 coding sequence.
 XX
 KW Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
 XX cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatocellular;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..1260
 FT CDS
 FT /*tag= a
 FT /product= "protein kinase 13887"


```

FT      /+tag= a
FT      /tag= .1529
FT      /product= "protein kinase 13887"
FT      /note= "coding sequence specifically claimed"
FT      3'UTR
FT      /tag= c
FT      /tag= c
XX      MO200173050-A2.
XX      04-OCT-2001.
XX      23-MAR-2001; 2001WO-US009483.
XX      24-MAR-2000; 2000US-0191846P.
XX      (MILL-) MILLENNIUM PHARM INC.
XX      Meyers R;
XX      WPI; 2001-611632/70.
XX      P-PSDB; AAG65767.
XX      New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
XX      useful in diagnosis of cancer or cellular proliferation or
XX      differentiation disorders and to screen for polypeptide modulators useful
XX      to treat such conditions.
XX      Claim 1; Fig 10A-C; 16pp; English.
XX      The invention provides novel human protein kinase polypeptides, 3714,
XX      16742, 23546 and 13887 and nucleic acid molecules encoding them. The
XX      protein kinase polypeptides can be expressed by standard recombinant
XX      methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
XX      are useful for diagnostic and screening methods to identify subjects (at
XX      risk of) having cancer or cellular proliferation and/or differentiation
XX      disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
XX      modulators are useful for the treatment of cancer, particularly colon
XX      cancer or cellular proliferation and/or differentiation disorders. Other
XX      disorders associated with 3714, 16742, 23546 or 13887 expression or
XX      activity that can be treated include bone related disorders, inflammatory
XX      disorders, autoimmune diseases, cardiovascular disorders and liver
XX      diseases. The present sequence represents a human protein kinase
XX      polypeptide 13887 encoding cDNA
XX      Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other:
SQ      Query Match 99.7%; Score 1256.8; DB 4; Length 2622;
      Best Local Similarity 99.8%; Pred. No. 0;
      Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
      1 ATGGCGGAGATCCGCGCTGCGCTGGGCGGAGACCCGCGCTTTCTGAGGCGCTTCGGG 60
      267 ATGGCGGAGATCCGCGCTGCGCTGGGCGGAGACCCGCGCTTTCTGAGGCGCTTCGGG 326
      61 CGGCTGTGGAGTACAGAGCCGCTGTGGGTAGCGGCTCTCGCTCGGCTGTATTCGGGTT 120
      327 CGGCTGTGGAGTACAGAGCCGCTGTGGGTAGCGGCTCTCGCTCGGCTGTATTCGGGTT 386
      121 CGGCTGTGGAGTACAGAGCCGCTGTGGGTAGCGGCTCTCGCTCGGCTGTATTCGGGTT 180
      387 CGGCTGTGGAGTACAGAGCCGCTGTGGGTAGCGGCTCTCGCTCGGCTGTATTCGGGTT 446
      181 ACCACCGGGGCTGCGGCTCTGTGCGCGGATATGTTTCGCAAGAGAGGCGGCGCTG 240
      447 ACCACCGGGGCTGCGGCTCTGTGCGCGGATATGTTTCGCAAGAGAGGCGGCGCTG 506
      241 GAAACAGTTGAGGCTACAGAAACATGTGATCTTTGTATGAGTGTATTCATTCACCTTT 300
      507 GAAACAGTTGAGGCTACAGAAACATGTGATCTTTGTATGAGTGTATTCATTCACCTTT 566
      301 TGTCCAAATGTGCAATCAGCTGTCTGTGTTGATCTGATCTCTGATCTGATCTGTTTGGAA 360
      |||||||
  
```

```

Db      567 TCTCCAAATGTGCAATCAGCTGTCTGTGTTGAACTCTCGGATGTGAGTGTTCGAA 626
Qy      361 TTGCTCTTATATTCAGTACACAGGGTGTTCATGTGATGATACAGATTTGCCCGGA 420
Db      627 TTGCTCTTATATTCAGTACACAGGGTGTTCATGTGATGATACAGATTTGCCCGGA 686
Qy      421 GATGTTTTGAGGCGCTTGTCTTTCTTCATCATGAGGCGTATGTCCATGCGGACTCAAA 480
Db      687 GATGTTTTGAGGCGCTTGTCTTTCTTCATCATGAGGCGTATGTCCATGCGGACTCAAA 746
Qy      481 CCACGTAACATATTCGTGAGTGCAGAGATGAAATGTTTTTAACTGATGACTTTGACTT 540
Db      747 CCACGTAACATATTCGTGAGTGCAGAGATGAAATGTTTTTAACTGATGACTTTGACTT 806
Qy      541 ACCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGAGGATTCGGCTCCA 600
Db      807 AGCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGAGGATTCGGCTCCA 866
Qy      601 GAAGCAAAATTCGAAATTTGCTTGCCCGCAGGCTGCGCTGAGATGATACAGATGTACC 660
Db      867 GAAGCAAAATTCGAAATTTGCTTGCCCGCAGGCTGCGCTGAGATGATACAGATGTACC 926
Qy      661 TGAGCTGTGATCTGTGAGGCTAGGAATCATTTTAACTGAGAAATGTTCTCAGGAATGAA 720
Db      927 TGAGCTGTGATCTGTGAGGCTAGGAATCATTTTAACTGAGAAATGTTCTCAGGAATGAA 986
Qy      721 CTGAAACATACAGTACAGATCTCAGGAATGGAAGGCAACAGTTCTGTATTTATGATAC 780
Db      987 CTGAAACATACAGTACAGATCTCAGGAATGGAAGGCAACAGTTCTGTATTTATGATAC 1046
Qy      781 ATATTTGCCAGTAAGCAATGAGTGAATGCGCAATTCAGCTTATCAGCTTAAGACACTT 840
Db      1047 ATATTTGCCAGTAAGCAATGAGTGAATGCGCAATTCAGCTTATCAGCTTAAGACACTT 1106
Qy      841 ATCAAAAGCATCTTCATGATATCCAAAGCAAGAAATTCGCTGAAATGAGCAATTTGTC 900
Db      1107 ATCAAAAGCATCTTCATGATATCCAAAGCAAGAAATTCGCTGAAATGAGCAATTTGTC 1166
Qy      901 AGCCCATCTTTAGCATTCCTTTTCCCTCATATTTGAAGATCTGTATGCTTCCACT 960
Db      1167 AGCCCATCTTTAGCATTCCTTTTCCCTCATATTTGAAGATCTGTATGCTTCCACT 1226
Qy      961 CCAAGTCTAAGCTGCTGATGTGCTGATGATGATTTATCTTTGGAAATGAAAGAAATAT 1020
Db      1227 CCAAGTCTAAGCTGCTGATGTGCTGATGATGATTTATCTTTGGAAATGAAAGAAATAT 1286
Qy      1021 GAAGATGTTGTAGAAATGTAAGAAAGAGTCAAAAATATGACCACTGATCTCTA 1080
Db      1287 GAAGATGTTGTAGAAATGTAAGAAAGAGTCTCAAAAATATGACCACTGATCTCTA 1346
Qy      1081 CTTGTTCCAAAGAAATCTTGCGCAGAGACAAAGTCTTGTGTGATATGCAAAATGCTGCT 1140
Db      1347 CTTGTTCCAAAGAAATCTTGCGCAGAGACAAAGTCTTGTGTGATATGCAAAATGCTGCT 1406
Qy      1141 GATTCCAAAGCTGCGCAGAAATTTACTGACTGGAAGAGATTTGATGGAAGTTTGTGTG 1200
Db      1407 GATTCCAAAGCTGCGCAGAAATTTACTGACTGGAAGAGATTTGATGGAAGTTTGTGTG 1466
Qy      1201 GCTACATTTACCGGCTGAGTCTTACAAAGAGGGAATTCGTATCAAACTTGCTTTAA 1260
Db      1467 GCTACATTTACCGGCTGAGTCTTACAAAGAGGGAATTCGTATCAAACTTGCTTTAA 1526
  
```

RESULT 6
AAH72714
ID AAH72714 standard; cDNA; 4065 BP.

AAH72714;
AC AAH72714;
XX 19-SBP-2001 (first entry)
DT 19-SBP-2001 (first entry)
XX 19-SBP-2001 (first entry)
XX Human cervical cancer marker nucleic acid 3988.

KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX Homo sapiens.
 XX MO200142467-A2.
 XX
 PD 14-JUN-2001.
 PF 08-DEC-2000; 2000WO-US033312.
 XX
 PR 08-DEC-1999; 99US-0169681P.
 PR 21-DEC-1999; 99US-0171350P.
 PR 14-MAR-2000; 2000US-0189315P.
 PR 12-MAY-2000; 2000US-0203791P.
 PR 09-JUN-2000; 2000US-0210600P.
 PR 21-JUL-2000; 2000US-0220114P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Deeds J, Berger A, Zhao X;
 XX WPI; 2001-375006/39.
 DR
 XX
 PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 PT for assessing and detecting compounds for treating the cancer.
 XX
 PS Claim 1; Page 810-811; 1051pp; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy
 CC
 SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;
 Query Match 99.7%; Score 1256.8; DB 4; Length 4065;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGGCGGAGATCCGGCTGCGCTGCGGCGGCGGAGCGCGCGCTTTCGAGGCGCTTCGGG 60
 DB 268 ATGGCGGAGATCCGGCTGCGCTGCGGCGGCGGAGCGCGCGCTTTCGAGGCGCTTCGGG 327
 QY 61 CGGCTGTGGCAGGTACAGAGCCGCTGTGGTACGCGCTCCCTCCGCTCGGTATTCGGGTT 120
 DB 328 CGGCTGTGGCAGGTACAGAGCCGCTGTGGTACGCGCTCCCTCCGCTCGGTATTCGGGTT 387
 QY 121 CGGCTGTGGCAGACCTTGCTGCGCCCGCGCGCTTCAAGAGATTTTCCGCCAGGA 180
 DB 388 CGCTGTGGCGGCAACCTGTGCTGCGCCCGCGCGCTTCAAGAGATTTTCCGCCAGGA 447
 QY 181 ACCACCGGGGCTGCGGCTCTGCGCGGAGTATGTTTCCGAAAGAGAGGGGCGGCTG 240
 DB 448 ACCACCGGGGCTGCGGCTCTGCGCGGAGTATGTTTCCGAAAGAGAGGGGCGGCTG 507
 QY 241 GAACAGTTCCAGGCTCAGAGAAACATCTGACTTTGTATGAGTGTTCACATCACTTT 300
 DB 508 GAACAGTTCCAGGCTCAGAGAAACATCTGACTTTGTATGAGTGTTCACATCACTTT 567
 QY 301 TCTCCAAATGTGCGATCAGCGCTGTCTGTGCTTGAACCTCTGAGATGTCACTTTGGAA 360
 DB 568 TCTCCAAATGTGCGATCAGCGCTGTCTGTGCTTGAACCTCTGAGATGTCACTTTGGAA 627
 QY 361 TTGCTCTTATATTCAGATCAGCGGTTGTTCATGTGATGATGATAGATATGCGCCGA 420
 DB 628 TTGCTCTTATATTCAGATCAGCGGTTGTTCATGTGATGATGATAGATATGCGCCGA 687
 QY 421 GATGTTTGGAGCCCTTGCTTTTCTTCATCATGAGGCGTATGTCCATGCGGACCTCAAA 480

DB 688 GATGTTTGGAGCCCTTGCTTTCTTCATCATGAGGCGTATGTCCATGCGGACCTCAAA 747
 QY 481 CCAGCTAACATATTTGTGAGTSCAGAGATGATGATTTTAACTGATGACTTTGAACTT 540
 DB 748 CACGTAACATATTTGTGAGTSCAGAGATGATGATTTTAACTGATGACTTTGAACTT 807
 QY 541 AGCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGACGGGTATCGGCTCCA 600
 DB 808 AGCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGACGGGTATCGGCTCCA 867
 QY 601 GAAGCAGAAATTCAGAAATTTGCTGGCCAGAGTGGCTGCGAGAGTATCAGAAATGTC 660
 DB 868 GAAGCAGAAATTCAGAAATTTGCTGGCCAGAGTGGCTGCGAGAGTATCAGAAATGTC 927
 QY 661 TCAGCTGTGATCTGAGAGCTTGAAGATCATTTTATCTGAAATGTTCTCAGAAATGAA 720
 DB 928 TCAGCTGTGATCTGAGAGCTTGAAGATCATTTTATCTGAAATGTTCTCAGAAATGAA 987
 QY 721 CTGAAACATACAGTCAAGATCTCAGAGATGAGAGGCAACAGTTCTGTATTAATGATC 780
 DB 988 CTGAAACATACAGTCAAGATCTCAGAGATGAGAGGCAACAGTTCTGTATTAATGATC 1047
 QY 781 ATATTTGCCAGTAAAGCAGTGTGATGCGCAATTCAGCTTATCAGCTTAAGAGACTT 840
 DB 1048 ATATTTGCCAGTAAAGCAGTGTGATGCGCAATTCAGCTTATCAGCTTAAGAGACTT 1107
 QY 841 ATCAAAAGCATGCTCATGATGATCCAAAGCAGAAATTCCTGCTGAATGGCATTTGTC 900
 DB 1108 ATCAAAAGCATGCTCATGATGATCCAAAGCAGAAATTCCTGCTGAATGGCATTTGTC 1167
 QY 901 ACCCATCTTTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATCTGTATGCTTCCACT 960
 DB 1168 ACCCATCTTTTGAAGATTCCTTTGCCCCCTCATATTTGAAGATCTGTATGCTTCCACT 1227
 QY 961 CCAGTGTCAAGCTGTGATGATGCTGTGATGATGATTTCTTGGAAATGAGAGAAAT 1020
 DB 1228 CCAGTGTCAAGCTGTGATGATGCTGTGATGATGATTTCTTGGAAATGAGAGAAAT 1287
 QY 1021 GAAGATGTTGTAAGATGTAAGAGAGAGTGTCAAAATATGACCAAGTGTATCTGTA 1080
 DB 1288 GAAGATGTTGTAAGATGTAAGAGAGAGTGTCAAAATATGACCAAGTGTATCTGTA 1347
 QY 1081 CTTGTTCCAAAGAAATCTTGCGCAGAGACAAAGCTTTTGTGATGATGCAAAATGCTG 1140
 DB 1348 CTTGTTCCAAAGAAATCTTGCGCAGAGACAAAGCTTTTGTGATGATGCAAAATGCTG 1407
 QY 1141 GATTCCAAAGCTGCGCAGAAATTAAGTGAAGAGATGTTGATGGAGATTTGTTG 1200
 DB 1408 GATTCCAAAGCTGCGCAGAAATTAAGTGAAGAGATGTTGATGGAGATTTGTTG 1467
 QY 1201 GCTAATTTTACCGCTGAGTCCCTAACAAGAGGATATCTGATCAAACTTGTGCTTAA 1260
 DB 1468 GCTAATTTTACCGCTGAGTCCCTAACAAGAGGATATCTGATCAAACTTGTGCTTAA 1527
 RESULT 7
 ADL45238
 ID ADL45238 standard; DNA; 4065 BP.
 XX
 AC ADL45238;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 XX Human ovarian cancer DNA marker #19128.
 DE Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
 KW Homo sapiens.
 OS Homo sapiens.
 XX
 FN WO200170979-A2.
 XX
 PD 27-SEP-2001.
 XX

PF 21-MAR-2001; 2001WO-US009126.
 XX
 PR 21-MAR-2000; 2000US-0191031P.
 PR 25-MAY-2000; 2000US-0207214P.
 PR 15-JUN-2000; 2000US-0211940P.
 PR 07-JUL-2000; 2000US-0216820P.
 PR 25-JUL-2000; 2000US-0220661P.
 PR 21-DEC-2000; 2000US-0257672P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Lee J, Little J;
 DR MPI; 2001-611502/70.
 XX
 PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
 PT cancer cells as compared to their normal non-cancerous ovarian cells are
 PT used to characterize stage, grade, histological type of ovarian cancer.
 PS Disclosure; SEQ ID NO 19128; 106bp; English.
 XX
 CC The invention relates to nucleic acid markers which are overexpressed in
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-
 CC cancerous) ovarian cells. The invention also relates to polypeptides
 CC encoded by the markers, antibodies that selectively bind to the
 CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
 CC of developing ovarian cancer involving inhibiting expression of a gene
 CC corresponding to a marker of the invention and a method of treating a
 CC patient afflicted with ovarian cancer comprising providing to cells of
 CC the patient an antisense oligonucleotide complementary to a marker of the
 CC invention. The markers are useful for assessing if a patient is afflicted
 CC with ovarian cancer, which involves comparing the level of expression of
 CC a marker in a patient sample and a normal level of expression of the
 CC marker in a control non-ovarian cancer sample. A difference between the
 CC expression levels indicates ovarian cancer. The level of expression of a
 CC marker corresponds to a secreted protein or to a transcribed
 CC polynucleotide or its portion. The level of expression of the marker is
 CC assessed by detecting the presence in the sample, a protein or protein
 CC fragment corresponding to the marker. The presence of protein or protein
 CC fragment is detected using an antibody that specifically binds with the
 CC protein or protein fragment. Alternatively, the level of expression of
 CC the marker is assessed by detecting the presence of a transcribed
 CC polynucleotide which anneals with the marker or anneals with a portion of
 CC the polynucleotide comprising the marker, under stringent conditions. The
 CC marker is also used for monitoring the progression of ovarian cancer in a
 CC patient which involves detecting expression of the marker in a patient
 CC sample at a first point in time, repeating the method at a subsequent
 CC time and comparing the level of expression. The method is carried out
 CC using an ovarian tissue sample. A composition comprising a marker,
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.
 CC This sequence represents a human ovarian cancer DNA marker of the
 CC invention.
 XX
 SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;
 Query Match 99.7%; Score 1256.8; DB 5; Length 4065;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 448 ACCACCGGGGCTGCGGCTCTCTGCGCCGAGATAGTGTTCGCAAGAGAGGCGGCGCTG 507
 QY
 DB 241 GAACAGTTGACAGGGTACAGAAACATCGACCTTTGATGAGTGTTCACATCCACTTT 300
 DB 508 GAACAGTTGACAGGGTACAGAAACATCGACCTTTGATGAGTGTTCACATCCACTTT 567
 QY 301 TCTCCAAATGTGCAATCAAGCTGTCTGTGCTTGAATCTCTGATGTCAGTGTTCGAA 360
 DB 568 TCTCCAAATGTGCAATCAAGCTGTCTGTGCTTGAATCTCTGATGTCAGTGTTCGAA 627
 QY 361 TTGCTCTTATATTCAGTACCAAGGGTGTTCATGATGATATACAGATTCGCCCGCA 420
 DB 628 TTGCTCTTATATTCAGTACCAAGGGTGTTCATGATGATATACAGATTCGCCCGCA 687
 QY 421 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
 DB 688 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 747
 QY 481 CCACGTACATATTTGAGAGTCAGAGATGAAATGTTTAACTCATTTGACTT 540
 DB 748 CCACGTACATATTTGAGAGTCAGAGATGAAATGTTTAACTCATTTGACTT 807
 QY 541 AGCTTAAAGAAAGCAATCAGATGTAAATATTCAGACAGAGGGTATGCGGCTCCA 600
 DB 808 AGCTTAAAGAAAGCAATCAGATGTAAATATTCAGACAGAGGGTATGCGGCTCCA 867
 QY 601 GAAGCAGATTTGCAAAATGCTTGGCCGAGGCTGCGCTCAGAGTATCAGAAATGATCC 660
 DB 868 GAAGCAGATTTGCAAAATGCTTGGCCGAGGCTGCGCTCAGAGTATCAGAAATGATCC 927
 QY 661 TGAAGCTTGTATCTGTGAGAGCTTGAAGATCATTTTACTGAAATGTTCTCAGAAATGAA 720
 DB 928 TGAAGCTTGTATCTGTGAGAGCTTGAAGATCATTTTACTGAAATGTTCTCAGAAATGAA 987
 QY 721 CTGAAACATACATGATCAGATCTCAGAAATGAAAGCAAGTTTCGCTATTATTGATC 780
 DB 988 CTGAAACATACATGATCAGATCTCAGAAATGAAAGCAAGTTTCGCTATTATTGATC 1047
 QY 781 ATATTGTCAGATTAAGAGATGATGATGCGCAATTCAGCTTATCAGTAAGACCTT 840
 DB 1048 ATATTGTCAGATTAAGAGATGATGATGCGCAATTCAGCTTATCAGTAAGACCTT 1107
 QY 841 ATCAAAAGCATCTTCATGATGATTCAGAGCAGAAAGATTCCTGCTGAATGSCATTGTGC 900
 DB 1108 ATCAAAAGCATCTTCATGATGATTCAGAGCAGAAAGATTCCTGCTGAATGSCATTGTGC 1167
 QY 901 AGCCCATTTTTCATGATTCCTTTGCCCCCTCATTTTGAAGATCTGCTATGCTTCCACT 960
 DB 1168 AGCCCATTTTTCATGATTCCTTTGCCCCCTCATTTTGAAGATCTGCTATGCTTCCACT 1227
 QY 961 CCAGTGTAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 1228 CCAGTGTAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1287
 QY 1021 GAAGATGTTGTAAAGATGTAAGAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTA 1080
 DB 1288 GAAGATGTTGTAAAGATGTAAGAGAGATGTAAGAGATGTAAGAGATGTAAGAGATGTA 1347
 QY 1081 CTTGTTTCAAAAGAAATCTGCGCAGAGACAAAGCTTGTGTAATGCAAAATCTGCT 1140
 DB 1348 CTTGTTTCAAAAGAAATCTGCGCAGAGACAAAGCTTGTGTAATGCAAAATCTGCT 1407
 QY 1141 GATTCCAAAGCTGCGCAGAAATTTACTGACTGAGAGATGTTGATGAGGAAATTTGTG 1200
 DB 1408 GATTCCAAAGCTGCGCAGAAATTTACTGACTGAGAGATGTTGATGAGGAAATTTGTG 1467
 QY 1201 GCTAATTTTACCCGCTGAGTCTTCAAGAGGGATATCTGTATCAAACTTGTCTTTAA 1260
 DB 1468 GCTAATTTTACCCGCTGAGTCTTCAAGAGGGATATCTGTATCAAACTTGTCTTTAA 1527
 RESULT 8
 AAS06734

AA06734 standard; cDNA; 1260 BP.
AA06734;
12-SEP-2001 (first entry)
Polynucleotide sequence encoding human protein kinase #34.
Human; protein kinase; PK; STK; cancer; cardiovascular disease;
metabolic disorder; immune related disease; neurological disorder;
neurodegenerative disorder; inflammatory disorder; infectious disease;
reproductive disorder; gene therapy; ss.
Homo sapiens.
WO200138503-A2.
31-MAY-2001.
22-NOV-2000; 2000WO-US032085.
24-NOV-1999; 99US-0167482P.
(SUGEN) SUGEN INC.
PLOWMAN GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
Flanagan P, Clary D;
MPI; 2001-343950/36.
P-SDB; AAU03534.
Nucleic acids encoding human kinase polypeptides, useful for preventing
diagnosing and/or treating e.g. cancer, immune, cardiovascular and
neural-associated diseases, and microbial infections.
Example 1; Fig 1; 433bp; English.
AA06701-AA06757 encode for novel human protein kinases #1-57. The novel
protein kinases have been identified as members of the tyrosine or
serine/threonine kinase (PTK and STK) families. The polynucleotides
encoding protein kinases and the polypeptides may be used in the
prevention, diagnosis and treatment of diseases associated with
inappropriate kinase expression. For example, they may be used to treat
cancers (especially cancers of hematopoietic origin), cardiovascular
disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
immune related diseases (e.g. rheumatoid arthritis), neurological
disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
disease (e.g. HIV) and reproductive disorders (e.g. infertility).
Additionally, polynucleotides encoding protein kinases may be used for
gene therapy and as DNA probes in diagnostic assays. The protein kinase
polypeptides may be used as antigens in the production of antibodies
against the protein kinases and in assays to identify modulators of
protein kinase expression and activity

Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other:
Query Match 99.6%; Score 1255.2; DB 4; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCGGATCCGCGCTGCGCTGGGGCGCGAGACCGCGGTTTCTGAGGCGCTTCGGG 60
DB 1 ATGGCGGATCCGCGCTGCGCTGGGGCGCGAGACCGCGGTTTCTGAGGCGCTTCGGG 60
QY CGGCTGTGGAGGATACAGAGCCGCTGTGGAGGCGCTCCCGCTCGAGTATCGAGTT 120
DB 61 CGGCTGTGGAGGATACAGAGCCGCTGTGGAGGCGCTCCCGCTCGAGTATCGAGTT 120
QY CGGCTGTGGAGGATACAGAGCCGCTGTGGAGGCGCTCCCGCTCGAGTATCGAGTT 120
DB 121 CGGCTGTGGAGGATACAGAGCCGCTGTGGAGGCGCTCCCGCTCGAGTATCGAGTT 120
QY CGGCTGTGGAGGATACAGAGCCGCTGTGGAGGCGCTCCCGCTCGAGTATCGAGTT 120
DB 121 CGGCTGTGGAGGATACAGAGCCGCTGTGGAGGCGCTCCCGCTCGAGTATCGAGTT 120

QY 181 ACCACCGGGGCTGGGAGCCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGGCTG 240
DB 181 ACCACCGGGGCTGGGAGCCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
QY 241 GAACAGTTGACAGGATACAGAAACATGCTGATTTGTATGAGTGTTCATTCACCTT 300
DB 241 GAACAGTTGACAGGATACAGAAACATGCTGATTTGTATGAGTGTTCATTCACCTT 300
QY 301 TCTCCAAATGTCATACAGCTGTCTGTGCTTGAATCTTGGATGTCAGTGTTCGAA 360
DB 301 TCTCCAAATGTCATACAGCTGTCTGTGCTTGAATCTTGGATGTCAGTGTTCGAA 360
QY 361 TTGCTCTTATATTCAGTACAGGAGGTTTTCATGAGTATGATCAGATTCGCGCGA 420
DB 361 TTGCTCTTATATTCAGTACAGGAGGTTTTCATGAGTATGATCAGATTCGCGCGA 420
QY 421 GATGTTTGGAGGCTGCTTCTTCTATCATGAGGCTATGTCATGCGGACCTCAAA 480
DB 421 GATGTTTGGAGGCTGCTTCTTCTATCATGAGGCTATGTCATGCGGACCTCAAA 480
QY 481 CCACGTAACATTTGTGAGTGCAGAGATGAATGTTTAACTCATTTGATCTTGA 540
DB 481 CCACGTAACATTTGTGAGTGCAGAGATGAATGTTTAACTCATTTGATCTTGA 540
QY 541 AGCTTCAAAAGGAGCAATCAGATGTAAAGTATATTCAGACAGAGGATCGGCTCCA 600
DB 541 AGCTTCAAAAGGAGCAATCAGATGTAAAGTATATTCAGACAGAGGATCGGCTCCA 600
QY 601 GAAGCAGATTGCAAAATTTGCTTGGCCAGGCTGCGCTGAGTATGAGATGTA 660
DB 601 GAAGCAGATTGCAAAATTTGCTTGGCCAGGCTGCGCTGAGTATGAGATGTA 660
QY 661 TGAAGTGTGATGCTGAGAGCTGAGATCATTTTACTGAAATGTTCTCAGAAATGAA 720
DB 661 TGAAGTGTGATGCTGAGAGCTGAGATCATTTTACTGAAATGTTCTCAGAAATGAA 720
QY 721 CTGAAACATACAGTATCTCAGAAATGAGAGGCAACAGTCTGATTTATGATC 780
DB 721 CTGAAACATACAGTATCTCAGAAATGAGAGGCAACAGTCTGATTTATGATC 780
QY 781 ATATTTGCGAGTAAAGCAGTGTGATGCGCAATTTCCAGCTTACCTTAAAGACCTT 840
DB 781 ATATTTGCGAGTAAAGCAGTGTGATGCGCAATTTCCAGCTTACCTTAAAGACCTT 840
QY 841 ATCAAAAGCATCTTCATGATGATCCAGCAAGAAATCTGCTGAATGCAATTTGTC 900
DB 841 ATCAAAAGCATCTTCATGATGATCCAGCAAGAAATCTGCTGAATGCAATTTGTC 900
QY 901 AGCCATTTCTTATGATCTTCTGCGCTGATTTGAAAGATCTGCTGCTCCACT 960
DB 901 AGCCATTTCTTATGATCTTCTGCGCTGATTTGAAAGATCTGCTGCTCCACT 960
QY 961 CCAAGTCTAAGCTGCTGATGCTGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 CCAAGTCTAAGCTGCTGATGCTGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 GAAATGTTGTAAGATGTAAGAGAGAGTCAAAATATGAGACCAAGTGTATCTTA 1080
DB 1021 GAAATGTTGTAAGATGTAAGAGAGAGTCAAAATATGAGACCAAGTGTATCTTA 1080
QY 1081 CTTGTTCCAAAGAAATCTGCGAGAGCAAGCTTTGTTGAGTATCAATGCTG 1140
DB 1081 CTTGTTCCAAAGAAATCTGCGAGAGCAAGCTTTGTTGAGTATCAATGCTG 1140
QY 1141 GATTCAAAGCTGCGAGAAATTAATGATGAGAGATGTTTATGAGAGTTTGTG 1200
DB 1141 GATTCAAAGCTGCGAGAAATTAATGATGAGAGATGTTTATGAGAGTTTGTG 1200
QY 1201 GCTACATTTACCGCTGAGTGCCTAAGAGAGGATATCTGTATCAAACTTGTCTTAA 1260
DB 1201 GCTACATTTACCGCTGAGTGCCTAAGAGAGGATATCTGTATCAAACTTGTCTTAA 1260

```

RESULT 9
AADI8818
ID AADI8818 standard; cDNA; 1824 BP.
XX
AC AADI8818;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human kinase (PKIN) -3 cDNA.
XX
KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
KW myasthenia gravis; cataracts; growth and development disorder;
KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW asthma; obesity; restorative therapy; cytosolic; immunomodulatory;
KW antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS /tag= a
FT CDS /product= "Human PKIN-3 protein"
XX
PN MO2001B1555-A2.
XX
XX 01-NOV-2001.
XX
PF 20-APR-2001; 2001MO-US012992.
XX
PR 20-APR-2000; 2000US-0199021P.
PR 28-APR-2000; 2000US-0200226P.
PR 05-MAY-2000; 2000US-0202339P.
PR 11-MAY-2000; 2000US-0203505P.
PR 18-MAY-2000; 2000US-0205564P.
PR 26-MAY-2000; 2000US-0207739P.
PR 01-JUN-2000; 2000US-0208795P.
XX
PI (INCY-) INCYTE GENOMICS INC.
PI
PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB,
PI Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG,
PI Paterson C, Burtill JD, Marcus GA, Zingler KA, Reipon SA, Lu Y,
PI Policky JL, Thornton M, Tang XT, Hafalia A, Elliott VS, Baughn MR,
PI Walsh RT, Ramkumar J, Borowsky ML, Au-Young J, Hillman JL,
PI Gururajan R;
XX
XX WPI; 2001-611740/70.
XX
XX P-PSDB; AAE11769.
XX
XX Human kinases and nucleic acids, useful for preventing diagnosing and
XX treating cancers, inflammation and immune disorders.
XX
XX Claim 5; Page 152; 166pp; English.
XX
XX The present invention relates to human kinases (PKIN) and the nucleic
XX acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
XX used in the prevention, diagnosis and treatment of diseases cancers,
XX adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
XX acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
XX gout, microbial infections, cardiovascular disease and/or inflammation,
XX myasthenia gravis, atherosclerosis, cataracts, osteoporosis, myocardial
XX infarction, cataract, growth and development disorder, seizure disorder,
XX pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
XX disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
XX PKIN may be used to treat disorders associated with decreased PKIN
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of PKIN by expressing inactive proteins or to
XX supplement the patient's own production of PKIN. PKIN nucleic acids may be
XX used to produce the PKIN polypeptide, by inserting the nucleic acids into

```

```

CC a host cell and culturing the cell to express the protein. PKIN nucleic
CC acid and its complementary sequences may also be used as DNA probes in
CC diagnostic assays to detect and quantitate the presence of similar
CC nucleic acid sequences in samples and therefore which patients may be in
CC need of restorative therapy. The present sequence is human PKIN-3 cDNA
XX
XX SQ Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;
XX
XX Query Match 99.6%; Score 1255.2; DB 4; Length 1824;
XX Best local similarity 99.8%; Pred. No. 0;
XX Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
OY 1 ATGCGCGGATCCGCTGCGCTGCGGCGGAGCCGCGGCTTTCTGAGGCTTCGCG 60
DB ATGCGCGGATCCGCGCTGCGCTGCGGCGGAGCCGCGGCTTTCTGAGGCTTCGCG 228
OY CGGCTGTGCAAGTACAGAGCCGCTCTGGGTACCGGCTCTCGGCTGTGATGGGTT 120
DB CGGCTGTGCAAGTACAGAGCCGCTCTGGGTACCGGCTCTCGGCTGTGATGGGTT 288
OY 121 CGCTGTGCGCAACCTTGCTGCGCCCGCGGCGCTTCAAGCATTTTGGCGCAGGA 180
DB 289 CGCTGTGCGCAACCTTGCTGCGCCCGCGGCGCTTCAAGCATTTTGGCGCAGGA 348
OY 181 ACCACCGGGGCTGCGGCTCTGCGCCCGAGATGATGTTTCCGAAAGAGGCGCGCTG 240
DB 349 ACCACCGGGGCTGCGGCTCTGCGCCCGAGATGATGTTTCCGAAAGAGGCGCGCTG 408
OY 241 GAACGTTGCAAGGTACAGAAACATCGACTTTGTATGAGGTATTAATCACTTT 300
DB 409 GAACGTTGCAAGGTACAGAAACATCGACTTTGTATGAGGTATTAATCACTTT 468
OY 301 TCTCCAAATGTCATCAAGCTGTCTGTGCTTGAATCTGATGATGATGTTTCGGA 360
DB 469 TCTCCAAATGTCATCAAGCTGTCTGTGCTTGAATCTGATGATGATGTTTCGGA 528
OY 361 TTGCTTTATTTTCAAGTACACAGGGTGTTCATGTGATGATATACAGATTGCCCGGA 420
DB 529 TTGCTTTATTTTCAAGTACACAGGGTGTTCATGTGATGATATACAGATTGCCCGGA 588
OY 421 GATGTTTGGAGGCGCTTCTTCTCATATGAGGGCTATGTCATGCGGACCTCAAA 480
DB 589 GATGTTTGGAGGCGCTTCTTCTCATATGAGGGCTATGTCATGCGGACCTCAAA 648
OY 481 CCACGTAAATTTTGGAGTGCAGAGATGATGATTTTAACTGATGATGATGATGAT 540
DB 649 CCACGTAAATTTTGGAGTGCAGAGATGATGATTTTAACTGATGATGATGATGAT 708
OY 541 AGCTTCAAAAGGCAATGATGATTAAGTATTTTCAAGCAGAGCGGATGCGGCTCCA 600
DB 709 AGCTTCAAAAGGCAATGATGATTAAGTATTTTCAAGCAGAGCGGATGCGGCTCCA 768
OY 601 GAAGCAGATTTGCAAAATGCTTGGCCGAGCGCTGCGGCTGAGATGATGATGATGAT 660
DB 769 GAAGCAGATTTGCAAAATGCTTGGCCGAGCGCTGCGGCTGAGATGATGATGATGAT 828
OY 661 TCAGCTTTGATCTGTGAGCTAGAGATCATTTTACTGGAATGTTTTCAGAAATGAAA 720
DB 829 TCAGCTTTGATCTGTGAGCTAGAGATCATTTTACTGGAATGTTTTCAGAAATGAAA 888
OY 721 CTGAAACATACATGATCTGATCTGAGATGAGAGCAAGTTCTGATTTATTTGATC 780
DB 889 CTGAAACATACATGATCTGATCTGAGATGAGAGCAAGTTCTGATTTATTTGATC 948
OY 781 AATTTTCCAGTAAGGAGTGTGTAATGCGGAATTTCCAGCTTACACTTAAGAGACCTT 840
DB 949 AATTTTCCAGTAAGGAGTGTGTAATGCGGAATTTCCAGCTTACACTTAAGAGACCTT 1008
OY 841 ATCAAAAGCATCTTCATGATGATCCAAAGCAGAAATTTCTGTGAAATGCAATTTGTC 900
DB 1009 ATCAAAAGCATCTTCATGATGATCCAAAGCAGAAATTTCTGTGAAATGCAATTTGTC 1068
OY 901 AGCCCATTTTGAAGATTCCTTTTGGCCCTCATATTAAGATGATGATGATGATGAT 960

```

```

Db      |||
1069  AGCCCAATCTTAGACATCTTTGGCCCTCAATATGAAGATCTGTCATGCTCCCACT 1128
Qy      961  CCAGTCTAAGACTGCTGATATGCTGATGATGATTAATCTTGGAAATGAAGGAATAT 1020
Db      1129  CCAGTCTAAGACTGCTGATGATGATGATGATTAATCTTGGAAATGAAGGAATAT 1188
Qy      1021  GAAGATGTTGTAAGATGTAAGAGAGATGTCANAAAAATATGACACAGTGTATCTTA 1080
Db      1189  GAAGATGTTGTAAGATGTAAGAGAGATGTCANAAAAATATGACACAGTGTATCTTA 1248
Qy      1081  CTTGTTCCAAAGAAATCTGCGCAGAGACAAGTCTTTGTTGATGATGCAATGCTGTG 1140
Db      1249  CTTGTTCCAAAGAAATCTGCGCAGAGACAAGTCTTTGTTGATGATGCAATGCTGTG 1308
Qy      1141  GATTCGAAGCTGCGCAGAAATTAATGATGATGAGAGATGTTGATGAGAAATTTGTTG 1200
Db      1309  GATTCGAAGCTGCGCAGAAATTAATGATGATGAGAGATGTTGATGAGAAATTTGTTG 1368
Qy      1201  GCTACATCTACCCGCTGAGTGCCTACAGAGGGGATCTGTATCAAACTTGCTTAA 1260
Db      1369  GCTACATCTACCCGCTGAGTGCCTACAGAGGGGATCTGTATCAAACTTGCTTAA 1428

```

RESULT 10
ADQ24401
ID ADQ24401 standard; DNA; 3998 BP.

ADQ24401;
26-AUG-2004 (first entry)

Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.

soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
ds.

Homo sapiens.

WO2004048938-A2.

10-JUN-2004.

26-NOV-2003; 2003WO-US038193.

26-NOV-2002; 2002US-0429739P.

(PROT-) PROTEIN DESIGN LABS INC.

Aziz N, Ginsburg WM, Zlotnick A;

WPI; 2004-441208/41.

Early detection of soft tissue sarcoma comprises determining expression
of a gene in a first soft tissue sample and a normal soft tissue sample
and comparing the gene expression, also useful in treating soft tissue
sarcoma.

Example 2; SEQ ID NO 7221; 210pp; English.

The invention relates to a novel method for detecting soft tissue sarcoma
which comprises obtaining a first soft tissue sample from an individual
and a normal soft tissue sample from the same or different individual,
determining the expression of a gene in both samples and comparing the
expression of the gene in both soft tissue samples, where a higher level
of protein expression in the first soft tissue sample indicates the
presence of soft tissue sarcoma. The method of the invention has
cytostatic applications and may be useful for detecting soft tissue
sarcoma, possibly via gene therapy or vaccine production. The nucleic
acid sequences may be useful in diagnostic and screening applications.
The current sequence is that of a human soft tissue sarcoma-upregulated
DNA of the invention. The current sequence is not shown within the
specification per se but was submitted in CD format by the inventor.

XX
SQ Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;
Query Match 93.3%; Score 1175; DB 12; Length 3998;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1215; Conservative 0; Mismatches 45; Indels 3; Gaps 1;

```

Qy      1  ATGGGGGATCCCGCTGCGCCTGCGGCGCGAGCCGCGCTTCTGAGGCTTCGCG 60
Db      160  ATGGCGGATCCCGCTGCGCCTGCGGCGCGAGCCGCGCTTCTGAGGCTTCGCG 219
Qy      61  CGGCTGCGCAGTACAGAGCCGCTGCGGCTTCGCGCTTCGCGCTGATGCGGTT 120
Db      220  CGGCTGCGCAGTACAGAGCCGCTGCGGCTTCGCGCTTCGCGCTGATGCGGTT 279
Qy      121  CGGCTGCGCAGTACAGAGCCGCTGCGGCTTCGCGCTTCGCGCTGATGCGGTT 180
Db      280  CGGCTGCGCAGTACAGAGCCGCTGCGGCTTCGCGCTTCGCGCTGATGCGGTT 339
Qy      181  ACCACCGGGGCTGCGGCTTCGCGCAGTATGATTCGCGCAAGAGAGGCGCGCTG 240
Db      340  ACCACCGGGGCTGCGGCTTCGCGCAGTATGATTCGCGCAAGAGAGGCGCGCTG 399
Qy      241  GAACAGTTGCAAGGTCACAGAAACATCGTACCTTGTATGAGTGTTCACATCCATT 300
Db      400  AAGTCTGTGTGACCTTAACCGAGAGTGAATTTGTATGAGTGTTCACATCCAGTT 459
Qy      301  TCT---CCAAATGTCCATCAGCTGTCTGTGCTTGAACCTCTGAGATGATGTTTCG 357
Db      460  GTTCTCCAAAGTGTCCATCAGCTGTCTGTGCTTGAACCTCTGAGATGATGTTTCG 519
Qy      358  GAATGCTCTTATATTCAGTACAGGAGTGTTCATGATGATGATGATGATGATGATG 417
Db      520  GAATGCTCTTATATTCAGTACAGGAGTGTTCATGATGATGATGATGATGATGATG 579
Qy      418  CGAGATGTTTGGAGCCCTTCTTCTTCATCAGAGGCTATGTCATGCGGACCTC 477
Db      580  CGAGATGTTTGGAGCCCTTCTTCTTCATCAGAGGCTATGTCATGCGGACCTC 639
Qy      478  AAACCACTGATATTTGTGAGTGCAGAGATGAATGTTTAACTCATTTGACTTTGA 537
Db      640  AAACCACTGATATTTGTGAGTGCAGAGATGAATGTTTAACTCATTTGACTTTGA 699
Qy      538  CTTAGCTTCAAAAGAGGCAATCAGAGTGAATGAATATTCAGACAGCGGATGCGCT 597
Db      700  CTTAGCTTCAAAAGAGGCAATCAGAGTGAATGAATATTCAGACAGCGGATGCGCT 759
Qy      598  CCAGAGCAGATTTGCAAAATTTGCTTGGCCCAAGGCTGCGCAGAGTATCAGAAAT 657
Db      760  CCAGAGCAGATTTGCAAAATTTGCTTGGCCCAAGGCTGCGCAGAGTATCAGAAAT 819
Qy      658  ACTCAGCTGTTGATCTGTGAGAGCTTAGAATCATTTTAACTGGAATGTTCTCAGAAT 717
Db      820  ACTCAGCTGTTGATCTGTGAGAGCTTAGAATCATTTTAACTGGAATGTTCTCAGAAT 879
Qy      718  AAACGAAACATACATGATCAGATCTCAGAAATGAAGGCAACAGTTCGTATTTAT 777
Db      880  AAACGAAACATACATGATCAGATCTCAGAAATGAAGGCAACAGTTCGTATTTAT 939
Qy      778  CACATATTTGCGAGTAAAGAGTGTGATGCGCAATTTCCAGCTTACCTTAAGAGC 837
Db      940  CACATATTTGCGAGTAAAGAGTGTGATGCGCAATTTCCAGCTTACCTTAAGAGC 999
Qy      838  CTTATCAAAAGATGCTTCATATGATTCAGACAGAAATTTCTGCTGAAATGCAATG 897
Db      1000  CTTATCAAAAGATGCTTCATATGATTCAGACAGAAATTTCTGCTGAAATGCAATG 1059
Qy      898  TGCAGCCATTTTATGATGCTTTTGGCCCTCATATTTGAAGATGATGATGCTGCC 957
Db      1060  TGCAGCCATTTTATGATGCTTTTGGCCCTCATATTTGAAGATGATGATGCTGCC 1119
Qy      958  ACTCAGTCTAAGACTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1017

```



```
QY 721 CTGAAACATACAGTCAGATCTCAGAAATGGAAGCAACAGTTCTGTATTAATGATC 780
DB 882 CTGAAACATACAGTCAGATCTCAGAGTGAAGGCAACAGTTCTGTATTAATGATC 941
QY 781 AATATTGGCCAGTAAGCAGTGGGTAATGCGCAATTCAGCCATATCAGCTTAAGAGACTT 840
DB 942 AATATTGGCCAGTAAGCAGTGGGTAATGCGCAATTCAGCCATATCAGCTTAAGAGACTT 1001
QY 841 ATCAAAAGCATGCTTCATGATGATTCACAGACAGAAATTCCTGCTGAAAATGGCATTTGTC 900
DB 1002 ATCAAAAGCATGCTTCATGATGATTCACAGACAGAAATTCCTGCTGAAAATGGCATTTGTC 1061
QY 901 AGCCCATTTCTTACATTCCTTTTGGCCCTCATATTTGAAGATTCGTCTGCTCCACT 960
DB 1062 AGCCCATTTCTTACATTCCTTTTGGCCCTCATATTTGAAGATTCGTCTGCTCCACT 1121
QY 961 CCAGTGTCAAGATGCTGTAATGCTGTAATGATGATTAATCTTGGGAATGAAGAGAAAT 1020
DB 1122 CCAGTGTCAAGATGCTGTAATGCTGTAATGATGATTAATCTTGGGAATGAAGAGAAAT 1181
QY 1021 GAAGATGTTGTAAGATGTAAGAGAGATGTCACAAAATATGACCAAGTGTATCTCTA 1080
DB 1182 GAAGATGTTGTAAGATGTAAGAGAGATGTCACAAAATATGACCAAGTGTATCTCTG 1241
QY 1081 CTGTGTTCCAAAGGAAATCTCTGCGACAGACAAAGTCTTTGTTGATGCAAAATGCTGCT 1140
DB 1242 CTGTGTTCCAAAGGAAATCTCTGCGACAGACAAAGTCTTTGTTGATGCAAAATGCTGCT 1301
QY 1141 GATTCGAAGCTGCGGAGAAATTAAGTCACTGGAAGATGTTGATGGAAGTGTGTTG 1200
DB 1302 GATTCGAAGCTGCGGAGAAATTAAGTCACTGGAAGATGTTGATGGAAGTGTGTTG 1361
QY 1201 GCTACATTTCTACCCGCTGAGTGCCTACAGAGGGAGATCTGTATCAAACTTGTCTTAA 1260
DB 1362 GCTACATTTCTACCCGCTGAGTGCCTACAGAGGGAGATCTGTATCAAACTTGTCTTAA 1421
```

RESULT 12

AAH70764

ID AAH70764 standard; cDNA; 529 BP.

AAH70764;

19-SEP-2001 (first entry)

Human cervical cancer marker nucleic acid 2038.

Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

Homo sapiens.

MO200142467-A2.

14-JUN-2001.

08-DEC-2000; 2000MO-US033312.

08-DEC-1999; 99US-0169681P.

21-DEC-1999; 99US-0171350P.

14-MAR-2000; 2000US-0189315P.

12-MAY-2000; 2000US-0203791P.

09-JUN-2000; 2000US-0210600P.

21-JUL-2000; 2000US-0220114P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

New isolated nucleic acid for diagnosing and treating cervical cancer and for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 434; 1051bp; English.

PS The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

SQ Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Query Match 36.4%; Score 458.6; DB 4; Length 529;
Best Local Similarity 97.8%; Pred. No. 3.5e-125;
Matches 486; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

```
QY 654 ATGTACCTCAGCTGTTGATTCGTGAGCTGAGATCATTTTACTGAAATGTTCTCAG 713
DB 22 AGTACCTCAGCTGTTGATTCGTGAGCTGAGATCATTTTACTGAAATGTTCTCAG 81
QY 714 AATGAACTGAAACATATACATGATCTCAGGATGAAAGCAACAGTTCTGTATTAT 773
DB 82 AATGAACTGAAACATATACATGATCTCAGGATGAAAGCAACAGTTCTGTATTAT 141
QY 774 TGATCAATATTGGCCAGTAAGCAGTGTGAATGCCCAATTCAGGCTTACCTTAG 833
DB 142 TGATCAATATTGGCCAGTAAGCAGTGTGAATGCCCAATTCAGGCTTACCTTAG 201
QY 834 AGACCTTATCAAAAGCATGCTTCATGATGATCAAGACAGAAATTCCTGTAATGGC 893
DB 202 AGACCTTATCAAAAGCATGCTTCATGATGATCAAGACAGAAATTCCTGTAATGGC 261
QY 894 ATTGTGACAGCCCATTCCTTACGATTCCTTTGCCCTCATATTTGAAGATTCGTGATGCT 953
DB 262 ATTGTGACAGCCCATTCCTTACGATTCCTTTGCCCTCATATTTGAAGATTCGTGATGCT 321
QY 954 TCCCACTCCAGTGTGAAGATGCTGTAATGCTGATGATGATTAATCTTGGGAATGAAGA 1013
DB 322 TCCCACTCCAGTGTGAAGATGCTGTAATGCTGATGATGATTAATCTTGGGAATGAAGA 381
QY 1014 GGAATATGAGA-TGTTGTGAGAAGATGTAAGA-GGATGCAAAAATATGACCAAGT 1071
DB 382 GGAATATGAGAAGTGTGTGAGAAGATGTAAGAAGAGAGATGCAAAAATATGACCAAGT 441
QY 1072 GTATCTCTACTGTTCCAAAGGAAATCCTGCGACAGGACAAAGTCTTGTGATGATGCA 1131
DB 442 GTATCTCTACTGTTCCAAAGGAAATCCTGCGACAGGACAAAGTCTTGTGATGATGCA 501
QY 1132 AATGCTGTGATTCGAA 1148
DB 502 AAGGCTGGGGATTCAAA 518
```

RESULT 13

AAA02535

ID AAA02535 standard; cDNA; 722 BP.

AAA02535;

19-MAY-2000 (first entry)

Human colon cancer cell line polynucleotide sequence SBQ ID NO:2526.

Human; colon cancer; tumour; diagnosis; gene expression product; probe;

detection; cancerous state; metastasis; identification; breast cancer;

estrogen receptor-negative breast cancer; lung cancer; ss.

Homo sapiens.

WO958675-A2.

XX 18-NOV-1999.
 PD 13-MAY-1999; 99MO-US010602.
 PF 14-MAY-1998; 98US-0085426P.
 PR 15-MAY-1998; 98US-0085537P.
 PR 15-MAY-1998; 98US-0085696P.
 PR 21-OCT-1998; 98US-0105234P.
 PR 27-OCT-1998; 98US-0105877P.
 XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSE INC.
 XX Williams LT, Becobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 PI Lamson G, Drmanac R, Ckvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leishkowitz D, Kita D, Garcia V, Jones LM, Stachne-Crain B;
 XX WPI; 2000-126369/11.
 DR Polynucleotide library used to determine cancerous states of mammalian
 XX cells.
 PT Claim 1; Page 1020; 1097pp; English.
 PS AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
 CC libraries constructed from human colon cancer cell lines. The present
 CC invention also describes a method of detecting differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test
 CC sample derived from a cell suspected of being cancerous, where detection
 CC of the differentially expressed gene product is correlated with a
 CC cancerous state of the cell from which the test sample was derived. The
 CC polynucleotide sequences can be used in a method for detecting
 CC differentially expressed genes correlated with a cancerous state of a
 CC mammalian cell. The polynucleotides can also be used as probes for
 CC detecting and mapping related genes. They can be used in diagnosis and
 CC prognosis of diseases and disorders (e.g. identification of pre-
 CC metastatic or metastatic cancerous states, stages of cancer, or
 CC responsiveness of cancer to therapy). This is particularly for breast
 CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
 CC negative breast cancer, lung cancer, and colon cancer
 CC XX

Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

Query Match 35.4%; Score 445.8; DB 3; Length 722;
 Best Local Similarity 93.0%; Pred. No. 2.7e-121;
 Matches 581; Conservative 0; Mismatches 29; Indels 15; Gaps 11;

QY 454 GAGGGCTATGTCCTCATGCGGACCTCAACACGTAATATTGTGAGTGAGAGAA 513
 DB 1 GAGGGCTATGTCCTCATGCGGACCTCAACACGTAATATTGTGAGTGAGAGAA 60
 QY 514 TGTTTTAACTCTATTGACTTTGACTTCAAGAAAGGCATAGAGTGAAGTAT 573
 DB 61 TGTTTTAACTCTATTGACTTTGACTTCAAGAAAGGCATAGAGTGAAGTAT 120
 QY 574 ATTACAGACAGCGGATGCGGCTCAGAGAGCAATTTGCAAAATGCTTGCGCCGCT 633
 DB 121 ATTACAGACAGCGGATGCGGCTCAGAGAGCAATTTGCAAAATGCTTGCGCCGCT 179
 QY 634 GGCTTCAGAGTATACAGAAATGACTGTTGATCTGTGAGCCCTAGAGATCAT 693
 DB 180 GGCTTCAGAGTATACAGAAATGACTGTTGATCTGTGAGCCCTAGAGATCAT 239
 QY 694 TTACTGGAATGTTCTCAGAGAACTGAACATACAGTCAAGATCTCAGAGATGAAG 753
 DB 240 TTACTGGAATGTTCTCAGAGAACTGAACATACAGTCAAGATCTCAGAGATGAAG 299
 QY 754 GGAAGAG--TTTCGATATTA--TTGATCACAATTTGCCAGTAA--GCAAGTGTGAATGC 809
 DB 300 GGAAGAGTTTCTGATATTTGATCACTATTTGCCAGTAAAGCAANTGTGAATGC 359

QY 810 CGCAATTCAGGCTTATCACTTAAGAGACCTTATCAAAAGCATGCTTCATGATCC--A 867
 DB 360 CGCAATTCAGGCTTATCACTTAAGAGACCTTATCAAAAGCATGCTTCATGATCCCA 419
 QY 868 AGCAGAGAAATTCCTGCTGAATGGCATT-GTGCAGCCCATCTTTAGC-ATTCCCTTTG 925
 DB 420 GCAGAGAAATTCCTGCTGAATGGCATTGGGCAGCCCATCTTTAGCATTCCCTTTG 479
 QY 926 CCCCTCATATTAAGATGTCATGCTTCCACTCCAGT-GCTAAGACTGCTGAATGTG 984
 DB 480 CCCCTCATATTAAGATGTCATGCTTCCACTCCAGTGAAGACTGCTGAATGTG 539
 QY 985 CT-GGATGATGATTACTTGGGAATGAA--GAGGAATGAAGATGTTGTAGAAGATGT- 1040
 DB 540 CTGGGATGATGATTACTTGGGAATGAAAGAGATTAAGAGATGTTGTAGAAGATGT 599
 QY 1041 -AAAGAGAGATGTCAAAATATGG 1064
 DB 600 AAAGAGAGATGTCAAAATATTTG 624

RESULT 14
 AAH71966
 ID AAH71966 standard; cDNA; 470 BP.
 XX
 XX AAH71966;
 AC
 XX
 DT 19-SEP-2001 (first entry)
 XX
 DE Human cervical cancer marker nucleic acid 3240.
 XX
 KM Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200142467-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000MO-US033312.
 XX
 PR 08-DEC-1999; 99US-0169681P.
 PR 21-DEC-1999; 99US-0171350P.
 PR 14-MAR-2000; 2000US-0189315P.
 PR 12-MAY-2000; 2000US-0203791P.
 PR 09-JUN-2000; 2000US-0210600P.
 PR 21-JUL-2000; 2000US-0220114P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Deede J, Berger A, Zhao X;
 XX
 DR WPI; 2001-375006/39.
 XX
 PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 XX for assessing and detecting compounds for treating the cancer.
 XX
 PS Claim 1; Page 636; 1051pp; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy
 CC XX

Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;
 Query Match 33.9%; Score 426.6; DB 4; Length 470;
 Best Local Similarity 98.7%; Pred. No. 1e-115;

| | |
|----|---|
| XX | RESULT 15 |
| XX | AAH69101 |
| XX | ID AAH69101 standard; cDNA; 461 BP. |
| XX | AAH69101; |
| XX | AC |
| XX | AAH69101; |
| XX | 19-SEP-2001 (first entry) |
| XX | DT |
| XX | Human cervical cancer marker nucleic acid 375. |
| XX | DE |
| XX | Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss. |
| XX | XX |
| XX | XX |
| XX | XX |
| XX | OS Homo sapiens. |
| XX | MO200142467-A2. |
| XX | PN |
| XX | 14-JUN-2001. |
| XX | PD |
| XX | 14-JUN-2001. |
| XX | PF |
| XX | 08-DEC-2000; 2000WO-US033312. |
| XX | PR |
| XX | 08-DEC-1999; 99US-0169681P. |
| XX | PR |
| XX | 21-DEC-1999; 99US-0171350P. |
| XX | PR |
| XX | 14-MAR-2000; 2000US-0189315P. |
| XX | PR |
| XX | 12-MAY-2000; 2000US-0203791P. |
| XX | PR |
| XX | 09-JUN-2000; 2000US-0210600P. |
| XX | PR |
| XX | 21-JUL-2000; 2000US-0220114P. |
| XX | PA |
| XX | (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC. |
| XX | PI |
| XX | Schlegel R, Deeds J, Berger A, Zhao X; |
| XX | WPI; 2001-375006/39. |
| XX | DR |
| XX | WPI; 2001-375006/39. |
| XX | PT |
| XX | New isolated nucleic acid for diagnosing and treating cervical cancer and |
| XX | PT for assessing and detecting compounds for treating the cancer. |
| XX | XX |
| XX | Claim 1; Page 168-169; 1051pp; English. |
| XX | XX |

| | | | | |
|---------------------------|-------|---------------------|-----------|-------------|
| Query Match | 29.8% | Score 375.4; | DB 4; | Length 461; |
| Best Local Similarity | 96.2% | Pred. No. 1.6e-100; | | |
| Matches 428; Conservative | 0; | Mismatches 11; | Indels 6; | Gaps 4 |

Search completed: November 29, 2004, 11:45:24
Job time : 656.5 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 4563 Seconds
(without alignments)
4342.403 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206
Sequence: 1 MAGSGCAGACPPRFLFARG.....VATFPLSAVKRYLVQTL 419

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 segs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+pn.model -DEV=xlp
-Q/cg2.1/USPTO.spool.p/US10798532/runat_29112004_112452_15890/app_query.fasta_1.1166
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORF=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532_QCGN_1_1_6317@runat_29112004_112452_15890 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb_da: *
2: gb_hg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 2198 | 99.6 | 1260 | 6 | CQ715025 Sequence |
| 2 | 2198 | 99.6 | 1260 | 6 | AX262521 Sequence |
| 3 | 2198 | 99.6 | 2598 | 6 | AX262519 Sequence |
| 4 | 2198 | 99.6 | 2909 | 9 | HSAS36197 Homo sapi |

| | | | | | |
|----|--------|------|--------|----|-------------------|
| 5 | 2198 | 99.6 | 4065 | 6 | CQ412057 Sequence |
| 6 | 2198 | 99.6 | 4065 | 6 | AX188296 Sequence |
| 7 | 2191 | 99.3 | 1260 | 6 | AX165543 Sequence |
| 8 | 2191 | 99.3 | 1824 | 6 | AX680136 Sequence |
| 9 | 2186 | 99.1 | 1633 | 10 | RNKRISRNA |
| 10 | 2186 | 99.1 | 3244 | 10 | RNU70372 |
| 11 | 2185 | 99.0 | 2041 | 10 | BC058732 |
| 12 | 2177 | 98.7 | 110000 | 2 | AC129141_1 |
| 13 | 2160 | 97.9 | 1254 | 10 | NMPKTS |
| 14 | 1993 | 90.3 | 1367 | 9 | BC026046 |
| 15 | 1893.5 | 85.8 | 2940 | 9 | HS805327 |
| 16 | 1764.5 | 80.0 | 1157 | 5 | CR386837 |
| 17 | 1721 | 78.0 | 1135 | 9 | AK058195 |
| 18 | 1634 | 74.1 | 2041 | 9 | BC014917 |
| 19 | 1050 | 47.6 | 259562 | 2 | AC094499 |
| 20 | 1045 | 47.4 | 601 | 10 | NMKISRNA |
| 21 | 908 | 41.2 | 105902 | 9 | AL359699 |
| 22 | 908 | 41.2 | 167995 | 2 | AC040901 |
| 23 | 904 | 41.0 | 211812 | 2 | AC027205 |
| 24 | 805.5 | 36.5 | 211635 | 2 | AC112763 |
| 25 | 789 | 35.8 | 529 | 6 | AX186343 |
| 26 | 766.5 | 34.7 | 722 | 6 | BD221413 |
| 27 | 729 | 33.0 | 470 | 6 | AX187546 |
| 28 | 724 | 32.8 | 100207 | 10 | AY180177 |
| 29 | 697.5 | 31.6 | 200095 | 10 | AC121606 |
| 30 | 697.5 | 31.6 | 220806 | 10 | AC123650 |
| 31 | 663 | 30.1 | 462 | 6 | AX184680 |
| 32 | 369 | 16.7 | 60640 | 10 | AY180176 |
| 33 | 369 | 16.7 | 185810 | 2 | AC136052 |
| 34 | 361 | 16.4 | 231 | 6 | BD220137 |
| 35 | 330.5 | 15.0 | 205752 | 2 | AC119431 |
| 36 | 311.5 | 14.1 | 419 | 6 | CQ393016 |
| 37 | 311.5 | 14.1 | 419 | 6 | CQ399411 |
| 38 | 267 | 12.1 | 185810 | 2 | AC136052 |
| 39 | 257 | 11.7 | 110000 | 8 | CR382128_04 |
| 40 | 252 | 11.4 | 1065 | 10 | HAMCD2A |
| 41 | 250.5 | 11.4 | 1269 | 10 | AB096078 |
| 42 | 250.5 | 11.4 | 1395 | 10 | AB096079 |
| 43 | 249.5 | 11.3 | 1400 | 10 | RATCDK2A |
| 44 | 249.5 | 11.3 | 2250 | 10 | BC061832 |
| 45 | 248.5 | 11.3 | 2809 | 3 | AY064464 |

ALIGNMENTS

RESULT 1
LOCUS CQ715025 1260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 959 from Patent WO02068579.
ACCESSION CQ715025
VERSION CQ715025.1 GI:42275882
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLES Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 959 06-SEP-2002;
PB Corporation (NY) (US)
FEATURES
source
1. 1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores: 2.81e-191 Length: 1260
Pred. No.: 2198.00 Matches: 418
Score:

Percent Similarity: 99.76%
 Best Local Similarity: 99.76%
 Query Match: 99.64%
 DB: 6
 Gaps: 0

Conservative: 0
 Mismatches: 1
 Indels: 0
 Gaps: 0

US-10-798-532-2 (1-419) x CQ715025 (1-1260)

QY 1 MetAlaGlySerGlyCybaLaTPGlyAlaGluProProArgPheLeuGluAlaPheGly 20
 Db 1 ATGGCGGGGATCCGGCTGGCGCTGGGGCGCGAGCCCGCGCTTTCTCGAGAGCGCTTCGGG 60
 QY 21 ArgLeuTPGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 61 CGGCTGTGGCAGGTACAGAGCCGCTCTGGGTAGCGGCTCCCTCCGCTCGGTATCCGGGTT 120
 QY 41 ArgCyCyGlyValnProGlySerProProGlyValAlaLeuGlyPheLeuProProGly 60
 Db 121 CGCTGTGGCGGCAACCTGTGCTGCCCCCGCGGCTCCAGAGCATTTGCGCGCAGGA 180
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyLeuArgAlaAlaLeu 80
 Db 181 ACCACCGGGGCTCGGCTCTGCGCGCGAGTATGTTCCGCAAGAGAGGGCGGCGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 241 GAACAGTTCAGGGGTACAGAAACATCGTGACTTGTATGAGAGTGTTCACATCCACTTT 300
 QY 101 SerProAsnValProSerArgCyLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
 Db 301 TCTCCAAATGTGCATCAGCTGTCTGTGCTTGAACCTCGAGATGTCAGTGTTCGGA 360
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCySerSerMetTrpMetIleGlnHisCybaLaArg 140
 Db 361 TTGCTCTTAAATTCAGAGTCCAGGGTGTTCATGTGATGATGATACAGCATTTGCGCCA 420
 QY 141 AspValLeuGlnValAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeu 160
 Db 421 GATGTTTGGAGGCCCTTGTCTTTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGlyCyPheLeuLeuLeuAspPheGlyLeu 180
 Db 481 CCACGTAACTATTTGGAGTGCAGAGATGAATGTTTTTAACTCATTTGAGCTT 540
 QY 181 SerPheLeuGlnGlyAsnGlnAspValTyrTyrIleGlnThrAspGlyTyrArgAlaPro 200
 Db 541 AGCTTCAGAGAGAGCATCAGATGTAAATATTCAGACAGACGGGTATCGGGCTCA 600
 QY 201 GluAlaGluLeuGlnAsnCyLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCyThr 220
 Db 601 GAAGCAGAAATTCGAAATGCTTGCCCAAGGCTGCGCTGCAGAGTATCAGAAATGTACC 660
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluLeuPheSerGlyMetLys 240
 Db 661 TCAAGCTGTGATCTGGAGACCTTAGAATATTTACTGAAATGTTCTCAGAGAAAGAAA 720
 QY 241 LeuLysHisThrValArgSerGlnGluTyrPheAlaAsnSerSerAlaIleIleAspHis 260
 Db 721 CTGAACATATCAGATCTCAGGAATGGAAGCAACAGTTCTGTATATTTAGTATCAC 780
 QY 261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
 Db 781 ATATTTCAGAGTAAAGAGGTGTGAATGCGCAATTCACGCTATACCTCAAGAGACCTT 840
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCyS 300
 Db 841 ATCAAAAGCATGTCTTCAATGATTCAGAGCAAGAAATTCCTGTGAAATGGCATTTGTC 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 901 AGCCCATTTCTTATGCAATCTTTTGGCCCTCATATTTGAATGAGTGGTCAATCTCCACT 960
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGlyGluTyr 340

Db 961 CCAGTGTCAAGCTGCTGATGTCTGTGATGATTTCTTGAGAAATGAGAGAAATAT 1020
 QY 341 GluAspValValGlnAspValLysGluGlnCyGlnLysTyrGlyProValValSerLeu 360
 Db 1021 GAAGATGTTGTAGAGATATTAAGAGAGAGTGTCAAAAATATGACCACTGATCTCA 1080
 QY 361 LeuValProLysGlnAsnProGlyValArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 Db 1081 CTGTGTCAAAGAGAAATCTGTGGCAGAGAGCAAGTCTTGTGTGAGATGCAATGCTGCT 1140
 QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyValArgMetPheAspGlyLysPheValVal 400
 Db 1141 GATTCAAAGCTGCGGAGAAATTAATCACTGAGAGAGATGTTGATGGGAAGTTGTGTG 1200
 QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 Db 1201 GCTACATTTACCCGCTGAGTGTCTTCAAGAGAGATATCTGTATCAAACTTGCTT 1257

RESULT 2
 AX262521 1260 bp DNA linear PAT 26-OCT-2001
 LOCUS AX262521
 DEFINITION Sequence 12 from Patent WO0173050.
 ACCESSION AX262521
 VERSION AX262521.1 GI:16511408
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 1 Meyers, R.
 3714, 16742, 23546, and 13887 novel protein kinase molecules and
 uses therefor
 Patent: WO 0173050-A 12 04-OCT-2001;
 Millennium Pharmaceuticals, Inc. (US)

FEATURES
 source location/Qualifiers
 1..1260
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 2,81e-191 Length: 1260
 Score: 2198.00 Matches: 418
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.64% Indels: 0
 DB: 6 Gaps: 0

US-10-798-532-2 (1-419) x AX262521 (1-1260)

QY 1 MetAlaGlySerGlyCybaLaTPGlyAlaGluProProArgPheLeuGluAlaPheGly 20
 Db 1 ATGGCGGGGATCCGGCTGGCGCTGGGGCGCGAGCCCGCGCTTTCTCGAGAGCGCTTCGGG 60
 QY 21 ArgLeuTPGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 61 CGGCTGTGGCAGGTACAGAGCCGCTCTGGGTAGCGGCTCCCTCCGCTCGGTATCCGGGTT 120
 QY 41 ArgCyCyGlyValnProGlySerProProGlyValAlaLeuGlyPheLeuProProGly 60
 Db 121 CGCTGTGGCGGCAACCTGTGCTGCCCCCGCGGCTCCAGAGCATTTGCGCGCAGGA 180
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyLeuArgAlaAlaLeu 80
 Db 181 ACCACCGGGGCTCGGCTCTGCGCGCGAGTATGTTCCGCAAGAGAGGGCGGCGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 241 GAACAGTTCAGGGGTACAGAAACATCGTGACTTGTATGAGAGTGTTCACATCCACTTT 300
 QY 101 SerProAsnValProSerArgCyLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120

[illegible]

```

QY 181 SerPheLeuGluGluValAsnGlnAaspValLysTyrIleGlnThrAaspGlyTyrArgAlapro 200
Db 808 AGCTTCAGAGAGAGCATCGATGTAAATATTATTCAGACAGCGGATTCGGCTCCCA 867
QY 201 GUAAGLULeuGlnAasnCybLeuAlaGlnAlaGlyLeuGlnSerAaspThrGluCysThr 220
Db 868 GAAGCAGAAATGGAAAATGCTTGGCCAGCGCTGGCTGAGATGATACAGAAATATAC 927
QY 221 SerAlValAaspLeuTrpSerLeuGlyIleIleLeuLeuGlnLysPheSerGlyMetLys 240
Db 928 TCGAGCTTGATGTGTGGAGCTTGAAGATCATTTTACTGGAATGTCTTCAGAAATGAAA 987
QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAasnSerSerAlaIleIleAaspHis 260
Db 988 CTGAAACATACATCGATCGATCTCAGAAATGGAAGCAACAGTTCTGCTATTATTGATCAC 1047
QY 261 IlePheAlaSerLysValAlaValAlaAlaIleProAlaTyrHisIleuAagaApleu 280
Db 1048 ATATTTGGCCAGTAAGCAGTGTGAATGCCGCAATTCAGCTTATCAGCTTAAGAGACTT 1107
QY 281 IleLysSerMetLeuHisAaspProSerArgArgIleProAlaGluMetAlaLeuCys 300
Db 1108 ATCAAAAAGCATGCTTCATGATGATCCAGACAGAAATTCCTGCTGAATGGCATTTGCTG 1167
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAaspLeuValMetLeuProThr 320
Db 1168 AGCCCATTTCTTGTAGCATTTCTTTGGCCCTCATATTGAAGATCTGGTTCATGCTTCCACT 1227
QY 321 ProValLeuArgLeuLeuAasnValIleuAaspAaspTyrIleuGlnGluGluGlyTyr 340
Db 1228 CCAAGTCTAAGACTGCTGAATGTGTGATGATGATTAATCTTAAGAATGAAGAGAAATAT 1287
QY 341 GluAaspValAlaGluAaspValLysGluGluCysGlnLysTyrGlyProValIleSerLeu 360
Db 1288 GAAGAGTGTGTAAAGATGTAAAGAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTA 1347
QY 361 LeuValProLysGluAasnProGlyArgGlyGlnValPheValGluTyrAlaAasnAlaGly 380
Db 1348 CTGTGTTCCAAAGGAAATCTCTGCGACAGACCAAGTCTTGTGATGATGCAAAATGCTGGT 1407
QY 381 AaspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAaspGlyAaspPheValIle 400
Db 1408 GATTCCAAAGCTGCGAGAAATATCTGACTGTGAAGATGTTTGAATGGAGATTTGTTGG 1467
QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db 1468 GCTACATTCTACCGCTGAGTGTCTAACAAGGGGATATCTGATCAAACTTGGCTT 1524

```

RESULT 4
 HSAS36197 2909 bp mRNA linear PRI 11-JAN-2003
 LOCUS HSAS36197
 DEFINITION Homo sapiens mRNA for KIS protein.
 ACCESSION AJ536197
 VERSION AJ536197.1 GI:27657360
 KEYWORDS KIS protein. (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 Bieche, I., Manceau, V., Curmi, P. A., Laurendon, I., Lachkar, S.,
 Leroy, K., Vaudaud, D., Sobel, A., and Manceau, A.
 Quantitative RT-PCR reveals a ubiquitous but preferentially neural
 expression of the KIS gene in rat and human
 Unpublished
 2 (bases 1 to 2909)
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1..2909
 /organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="RZPD:IMAGP998M213680"
/db_xref="taxon:9606"
/clone="I.M.A.G.E. Consortium Clonoid 145548"
/tissue_type="2 pooled tumors (clear cell type)"
1..2909
/gene="KIS"
/cds_start=1
/cds_end=1405
/gene="KIS"
/codon_start=1
/product="KIS protein"
/protein_id="CAD60192.1"
/db_xref="GI:27657361"
/translation="MAGSGCAGWAPPRFLAFLGLMVOQRSLGSSASVYRVRCG
NGSPGALKQFLPGTTGAASAAEYGRKRALIQLQGRNIVTLVYGVETIHSP
NVSRLLELFDIVSELVLSHQGSMNIIQCARVLEALFLHHEGVHADLK
PNIILMSAENEFCKIIDPLSPKENDQVRYIQDGYRAREALONCLAQGLSDTE
CTSAVDLMSLGIILEMFSKMLKRTYASQEKANSSAIIDHIFSKAVVNAATRAY
LNDLIKSMILHDSRRIPAEMLCSPPSIPAPHIEDLVMLPTVLLNLVLDYDL
ENBEYEDVEDVEKCEQKYGPVSLVLPKENPGRGVFEVYANAGDSKAQKLTGR
MDGKRVATFEPFLSAYKRGYLYQTL"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 9, 01e-191
Score: 2198.00
Percent Similarity: 99.76%
Best Local Similarity: 99.76%
Query Match: 9
DB: 9
Gaps: 0

```

US-10-798-532-2 (1-419) x HSAS36197 (1-2909)

```

QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
Db 146 ATGGCGGAGATCCGGCTGCGCCCTGGGGCGGAGCGCCGCTTTCTGGAGGCTTCCGG 205
QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db 206 CGGCTGTGGCAGGTACAGAGCGGTCTGGGTACCGGCTCCTCCGCTCGGTATCGGGTT 265
QY 41 ArgCysCysGlyAasnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
Db 266 CGCTGCTGGGCAACCTGTGCTGCGCCCGCGGCTTCAACAGATCTTGGCGCAGGA 325
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
Db 326 ACCACCGGGGCTGCGGCTCTGCGCCGAGATATGTTCCGGAAGAGAGGCGCGCTG 385
QY 81 GluGlnLeuGlnGlyHisArgAasnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db 386 GAACAGTTGCAAGGTCACAGAAACATCGCATTTGTATGAGTGTATCAATCCACTTT 445
QY 101 SerProAasnValProSerArgCysLeuLeuGluGluLeuAaspValSerValSerGlu 120
Db 446 TCTCCAAATGTCCATCAGCTGTCTGTGATGAATCTCTGATGTCACTGTTCGGGA 505
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db 506 TTGCTCTTATATTCCAGTACAGAGGTTCTTCATGTGATATACAGATTGTGCCGA 565
QY 141 AaspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAaspLeuLys 160
Db 566 GATGTTTGAAGGCTTGTCTTCTTCATCATGAGGCTAATGTCATGCGGACCTCAAA 625
QY 161 ProArgAasnIleLeuTrpSerAlaGluLeuGluCysPheLysLeuIleAaspPheGlyLeu 180
Db 626 CCAAGTAAATATTTGTGAGTGCAGAGATGAATGATTTAAACTATGACTTGGACTT 685
QY 181 SerPheLysGluGluValAsnGlnAaspValLysTyrIleGlnThrAaspGlyTyrArgAlapro 200
Db 686 AGCTTCAGAGAGCATCGATGTAAATATTATTCAGACAGACGCGATTCGGCTCCA 745

```


QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 DB 746 GAAGCAGATTGCAAAATTCCTTGCCCGAGCTGCGCTGCAGAGTACATACAGAAATGACC 805
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys 240
 DB 806 TCAGCTGTTGATCTGGAGCCTGAGAAATCATTTTACTGGAAATGTTTCCAGGAATGAAA 865
 QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
 DB 866 CTGAACCAATACAGTCCGATCTCAGGAATGAGAGCAACAGTTCCTCTATTAATGATCAC 925
 QY 261 IlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTyrHisIleuArgAspLeu 280
 DB 926 ATATTGGCCAGTAAACAGTGGTGAATGCCGCAATTCAGCCCTATCATCTAGAGACCTT 985
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 986 ATCAAAAGCATGCTTCATGATGATCCAGAGCAAGAAATTCCTGCAATGGCATTTGTC 1045
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 1046 AGCCCATTTCTTAAGCATTCCTTTTGCCCTCATATTGAAGATCTGGTCAATGCCACT 1105
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGly 340
 DB 1106 CCAAGTCTTAAGCTGCTGAATGCTGATGATGATTAATCTTGAGAAATGAGAGAAAT 1165
 QY 341 GluAspValAlaGluAspValLysGluGluCysGlnLysTyrGlyProValAlaSerLeu 360
 DB 1166 GAAGATGTTGTAAGACATGTAAGAGAGAGTCAAAATATAGACACAGTGGATCTCTA 1225
 QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 DB 1226 CTGTGTCGAAGAAATCTCTGCGCAGAGACAGATCTTTGTTGAGATGCAAAATGCTGG 1285
 QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValAla 400
 DB 1286 GATTCCAAAGCTGCGCAAAATTAATGACTGCAGAGAGATGTTGATGGAGATTGTTGG 1345
 QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 DB 1346 GCTACATTCACCCGCTGAGTGCCTACAGAGGGGATCTGTATCAAACTTGCTT 1402
 RESULT 5
 CQ412057 4065 bp DNA linear PAT 23-JAN-2004
 LOCUS Sequence 19128 from Patent WO0170979.
 DEFINITION CQ412057
 ACCESSION CQ412057
 VERSION CQ412057.1 GI:41119838
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 Lee,J. and Lillie,J.
 AUTHORS Genes, compositions, kits, and method for identification,
 TITLE assessment, prevention, and therapy of ovarian cancer
 JOURNAL Patent: WO 0170979-A 19128 27-SEP-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source Location/Qualifiers
 1..4065
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: 1.44e-190 Length: 4065
 Pred. No.: 2198.00 Matches: 418
 Score: 99.764 Conservative: 0
 Percent Similarity: 99.764 Mismatches: 1
 Best Local Similarity: 99.764

Query Match: 99.64% Indels: 0
 DB: 6 Gaps: 0
 US-10-798-532-2 (1-419) x CQ412057 (1-4065)
 QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProArgPheLeuGluAlaPheGly 20
 DB 268 ATGGGGGAGATCCGGCTGGCGCTGGGGCGGAGCGCGCGTTTCTGGAGGCTTCGGG 327
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerAlaSerValTyrArgVal 40
 DB 328 CGCTGTGGCAGGTACAGAGCCGCTGGGTAGCCGCTCTCCGCTCGGTATGCGGGT 387
 QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
 DB 388 CGCTGTGGGACCCCTGGCTGGCGCCCGCGGCTCAACAGATTCCTTGCGCGCAGGA 447
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaLeu 80
 DB 448 ACCACCGGGGCTGGCGCTGCGCCCGCAGTATGTTCCGCAAGAGAGGGCGGCGTG 507
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB 508 GAACAGTTGCAGGGTCACAGAAACATCGACTTTGTATGAGTGTTTAACAATCCACTT 567
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuAspValSerValSerGlu 120
 DB 568 TCTCCAAATGTCCATACGCTGTGTGTTGAATCTCGAGATGCAATGTTTCGGAA 627
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 628 TTGCTCTTATATTCAGTACACAGGGTGTTCATGATGATATACAGATTTGTCCGA 687
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisIleGluGlyTyrValHisAlaAspLeuLys 160
 DB 688 GATGTTTGGAGGCCCTTCTTCTCATCATGAGGGCTATGTCATGCGGACCTCAAA 747
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
 DB 748 CCAAGTACATATTTGAGAGTGCAGAGATGATATTTTAACTATGATGACTTGGACTT 807
 QY 181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
 DB 808 AGCTTCAAAGAGGAGCATCAGATGTAAGTATTCAGACAGAGGGATGGGCTCCA 867
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 DB 866 GAAGCAGAAATTCAGAAATTCCTTGCCCGAGCTGCGCTGCAGAGTACAGAAATGACC 927
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys 240
 DB 928 TCAGCTGTTGATCTTGAGCTGAGCTAGGAATCATTTTACAGAAATGTTTCCAGAAATGAAA 987
 QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
 DB 988 CTGAACCAATACAGTCCGATCTCAGGAATGAGAGCAACAGTTCCTCTATTAATGATCAC 1047
 QY 261 IlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTyrHisIleuArgAspLeu 280
 DB 1048 ATATTGGCCAGTAAACAGTGGTGAATGCCGCAATTCAGCCCTATCATCTAGAGACCTT 1107
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 1108 ATCAAAAGCATGCTTCATGATGATCCAGCAAGAAATTCCTGCAATGGCATTTGTC 1167
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 1168 AGCCCATTTCTTAAGCATTCCTTTTGCCCTCATATTGAAGATCTGGTCAATGCCACT 1227
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluGly 340
 DB 1228 CCAAGTCTTAAGACTGCTGATGATGATTAATCTTGAGAAATGAGAGAAATAT 1287

QY 341 GIUAAPVAlValGlUAAPVAllySGlUGlUCySGlnlySTyrglYProVAlValSerleu 360
DB 1288 GAAGATGTTGTAAGAATGTAAAGAGAGGTCTCAAAAATATGACAGAGGTATCTTA 1347
QY 361 LeuValProlYSGlUAsnProGlYArGlyGlnValPheValGluTYrAlaAsnAlaGly 380
DB 1348 CTGTGTCGAAAGGAAATCTGTGACAGAGCAAGTCTTGTGATGATGCAATGCTGTG 1407
QY 381 AAspSerlyAlaAlaGlnlySLeuLeuThrglYArGMeCPheAspGlyLySPhaVal 400
DB 1408 GATTCAGAAAGCTCGGAGAAATTAATTAAGTGAAGAGATGTTTGAAGAGATGTTGTG 1467
QY 401 AlaThrPheTyPProLeuSerAlaTYrlySArGlyTYrleuTYrGlnThrleu 419
DB 1468 GCTACATTCTACCGCTGAGTGTCTTACAGAGGGAGATCTGTATCAAACTTGCTT 1524

RESULT 6
AX188296 AX188296 4065 bp DNA linear PAT 06-AUG-2001
LOCUS Sequence 3991 from Patent WO0142467.
DEFINITION AX188296
ACCESSION AX188296
VERSION AX188296.1 GI:15139769
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Schlegel, R., Deede, J., Berger, A. and Zhao, X.
TITLE Gene, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 3991 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)

FEATURES
source Location/Qualifiers
1..4065
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1,44e-190 Length: 4065
Score: 2198.00 Matches: 418
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.64% Indels: 0
DB: Gaps: 0

US-10-798-532-2 (1-419) x AX188296 (1-4065)

QY 1 MetAlaGlySerGlyCySaLaATPGlYAlaGluProProArpPheLeuGlnAlaPheGly 20
DB 268 ATGGCGGGATCCGGCTGCGCTGGGGCGCGAGCCCGCGGTTTTCGAGGCTTCGGG 327
QY 21 ArGLeuTYrGlnValGlnSerArGLeuGlySerGlySerSerAlaSerValTYrArGVal 40
DB 328 CCGCTGTGCGAGGTACAGAGCCGTCTGGGTAGCGGCTCCCTCCGCGGTATCGGGTT 387
QY 41 ArGcyCySGlYAsnProGlYArGProProGlYAlaLeuLySLeuPheLeuProProGly 60
DB 388 CGCTGCTGGCGCAACCTGTGCTGCCCCCGCGGCTTCAAGCATTTTGGCGCAGGA 447
QY 61 ThrThGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArGlySGlUArGAlaAlaLeu 80
DB 448 ACCACCGGGGCTGGGCTCTGCGCGCGAGTATGTTTCGCAAAAGAGGGGGCGGTG 507
QY 81 GlUglnLeuGlnGlyhiSarGAsnIleValThrLeuTYrGlyValPheThrIlehiSph 100
DB 508 GAACAGTTGCGAGGTACAGAAACATCTGTGATGAGAGTGTTCACATCCACTTT 567
QY 101 SerProArValPProSerArGcySLeuLeuLeuGlnLeuAspValSerValSerGlu 120
DB 568 TCTCCAAATGTGCATCAGCTGTGTGTGCTTGAACCTCTGTGATGTCAGTGTTCGGA 627

QY 121 LeuLeuLeuTYrSerSerhiSGlnGlyCySerMetTTPMeCilleGlnhiCySaLaArG 140
DB 628 TTGCTCTTATATTCAGTACACAGGGTTGTTCTCAATGGATATACAGATTGTGCCCA 687
QY 141 AAspValLeuGlnAlaLeuAlaPheLeuThiSGlUGlUYrValhiAlaAspLeuLyS 160
DB 688 GATGTTTGGAGCCCTTGCTTTCTTCATCTAGAGGGCTATGTCATCGGACCTCAAA 747
QY 161 ProArGAsnIleLeuTYrSerAlaGluAsnGluCySPhaLeuIleAspPheGlyLeu 180
DB 748 CACGTAACATATGTGTGAGTCAAGATGATGTTTAACTCATTTGACTTTGACTT 807
QY 181 SerPheLySGlUlyAsnGlnAspVallySTyrlleGlnThrAspGlyTYrArGAlaPro 200
DB 808 AGCTTCAGAAAGGAGCATCAGATGTAAAGTATATTCAGACAGAGGGATGCGGCTCA 867
QY 201 GIUAAlaGluLeuGlnAsnCylLeuAlaGlnAlaGlyLeuGlnSerAspThrglUCySThr 220
DB 868 GAAGCAGAAATTGCAAAATTTGCTTGCCCGCGCTGCGACAGATGATCAGAAATGTACC 927
QY 221 SerAlaValAspLeuTYrSerLeuGlyIleleleleuGluMetPheSerGlyMetLyS 240
DB 928 TCAGCTGTGATCTGTGAGCTAGAGATCATTTTACGAAATGTTTCAGAAATGAAA 987
QY 241 LeuLySHeThrValArSGerGlnGluTYrlySAlaAsnSerSerAlaIleleleSph 260
DB 988 CTGAAGAACATACAGTCAAGATCTCAGAGATGGAAGCAAGTTCGTCTATTATTCATC 1047
QY 261 IlePheAlaSerlyAlaValAlaAsnAlaIleProAlaTYrhiSLeuArGLeu 280
DB 1048 ATATTTGCGCAGTAAGACATGTGTGATGCGCAATTCAGCTTACCTTAAGAACCTT 1107
QY 281 IleLySerMetLeuhiSAspAspProSerArGAlaGlyleProAlaGluMeAlaLeuCyS 300
DB 1108 ATCAAAAGCATCTTCATATATATCCAGACAGAAATTCCTGCTGAATGGCATTTGTC 1167
QY 301 SerProPhePheSerIleProPheAlaProhiSileGluAspLeuValMetLeuProThr 320
DB 1168 AGCCCATTTTACGATTCCTTTTGCCCTCATATTAAGATCTGTGATGCTCCACT 1227
QY 321 ProValLeuArGLeuLeuAsnValleuAspArpArTYrleuGlyAsnGluGluTYr 340
DB 1228 CCAGTCTTAAGCTGTGATGTCTGTGATGTATCTTGAGAAATGAAGAAATAT 1287
QY 341 GIUAAPVAlValGlUAAPVAllySGlUGlUCySGlnlySTyrglYProVAlValSerleu 360
DB 1288 GAAGATGTTGTAAGAATGTAAAGAGAGGTCTCAAAAATATGACAGAGGTATCTTA 1347
QY 361 LeuValProlYSGlUAsnProGlYArGlyGlnValPheValGluTYrAlaAsnAlaGly 380
DB 1348 CTGTGTCGAAAGGAAATCTGTGACAGAGCAAGTCTTGTGATGATGCAATGCTGTG 1407
QY 381 AAspSerlyAlaAlaGlnlySLeuLeuThrglYArGMeCPheAspGlyLySPhaVal 400
DB 1408 GATTCAGAAAGCTCGGAGAAATTAATTAAGTGAAGAGATGTTTGAAGAGATGTTGTG 1467
QY 401 AlaThrPheTyPProLeuSerAlaTYrlySArGlyTYrleuTYrGlnThrleu 419
DB 1468 GCTACATTCTACCGCTGAGTGTCTTACAGAGGGAGATCTGTATCAAACTTGCTT 1524

RESULT 7
AX166543 AX166543 1260 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 34 from Patent WO0138503.
DEFINITION AX166543
ACCESSION AX166543
VERSION AX166543.1 GI:14546888
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1

AUTHORS Ploewman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R.,
 Flanagan, P. and Clary, D.S.
 TITLE Novel human protein kinases and protein kinase-like enzymes
 JOURNAL Patent: WO 0138503-A 34 31-MAY-2001;
 Sugen, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..1260
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,236-190 Length: 1260
 Score: 2191.00 Matches: 417
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.32% Indels: 0
 Gaps: 0
 DB: 6
 US-10-798-532-2 (1-419) x AX166543 (1-1260)
 QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
 DB 1 ATGGCGGGATCCGGCTGCGCTGGGGCGCGAGCCCGCGGTTTCTGGAGGCTTCGGG 60
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB 61 CGGCTGTGGAGGTACAGAGCCGCTGTGGTACGGGCGCTCCCGCTCGGTATCCGGGTT 120
 QY 41 ArgCysCysGlyAsnProGlySerProProGlyValAlaLeuGlyAsnLeuProProGly 60
 DB 121 CGCTGTGGGGCAACCTGTGCTGCCCCCGCGCCCTCAAGCATCTTGGCCGCAAGGA 180
 QY 61 ThrThrGlyValAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
 DB 181 ACCACGGGGCTCGGCTCTGCGCGAGTATGTTCCGCAAGAGGGGGGGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAsnLeuValThrLeuTyrGlyValPheThrLeuHisPhe 100
 DB 241 GAACAGTTGCAGGGTCAACAGAAACATGTCATTTGATGAGTGTTCATTCACATCCACTT 300
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuAspValSerValSerGlu 120
 DB 301 TCTCCAAATGTGTCATCAGCTGTCTGTGCTTGAATCTCGATGTCAGTGTTCGGAA 360
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 361 TTGCTCTTATATTCAGTCCAGGCTGTTCATGTGATGATACAGCATTTGCCCCA 420
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuGly 160
 DB 421 GATGTTTGGAGGCCCTTGTCTTTCTTCATCATGAGGGCTATGTCATCGGAGCTCAAA 480
 QY 161 ProArgAsnLeuLeuTrpSerAlaGluAsnGluCysPheGlyLeuLeuAspPheGlyLeu 180
 DB 481 CCAGTAACTATATGTGAGTGTGAGAAATGTTTAACTCATGCTTGGACTT 540
 QY 181 SerPheGlySerGlnGlyAsnGlnAspValTyrTrpIleGlnThrAspGlyTyrArgAlaPro 200
 DB 541 AGCTTCAAGAAAGGCAATCGAGTGTAAATATATTCAGACAGCGGATTCGGGCTCCA 600
 QY 201 GluAlaGluLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 DB 601 GAAGCAGAAATTCGCTTGGCCAGGCTGCGCTGCAGAGTGTATACAGATGTAC 660
 QY 221 SerAlaValAspLeuTrpSerLeuGlyTyrIleLeuLeuGlnLeuMetPheSerGlyMetIys 240
 DB 661 TCAAGCTGTATCTGTGAGCCCTAGGATCATTTTACGTGAAATGTTCTCAGAAATGAAA 720
 QY 241 LeuLeuHisThrValArgSerGlnGluTyrPheAlaAsnSerSerAlaIleLeuAspHis 260
 DB 721 CTGAACATACAGTCAATCTCAGGATGAGGCAAAACGTTCTGTATATTATGATCAC 780

QY 261 IlePheAlaSerValAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuAspPhe 280
 DB 781 ATATTTCAGTAAGACATGTGTGATGCGCAATTCACGCTATCACCTAAGACCTT 840
 QY 281 IleYsSerMetLeuHisAspAspProSerArgGlyIleProAlaGluMetAlaLeuCys 300
 DB 841 ATCAAAAGCATCTTCATGATGATCCAGCAGAAAGATTCCTGCTGAATGCAATTGTGC 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 901 AGCCCATCTTATGATTCCTTTGCCCCCTCATATTAAGATCTGTGATCTCCACT 960
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTyrLeuGlyAsnGluGluGly 340
 DB 961 CCAGGCTAAGACTGTGATATGTCTGATGATGATTTATCTTGAGAAATGAGAAATAT 1020
 QY 341 GluAspValAlaGluAspValIysGluGluCysGlnTyrGlyProValValSerLeu 360
 DB 1021 GAAGATGTTGTGAAGATGTAAAGAGAGTCTCAAAAATATGACCACTGTATCTCA 1080
 QY 361 LeuValProIysGluAsnProGlyIysGlnValPheValGluTyrAlaAsnAlaGly 380
 DB 1081 CTGTTCCAAAGGAAATCTCTGCAGAGGACAAGCTTGTGATGATGCAATCTGCT 1140
 QY 381 AspSerValAlaAlaGlnIysLeuLeuThrGlyArgMetPheAspGlyIysPheValVal 400
 DB 1141 GATTCCAAAGCTGCGCAAAATTAATCTGACTGAGAGATTTGATGGAGAGTTGTGTG 1200
 QY 401 AlaThrPheTyrProLeuSerAlaTyrIysAspGlyTyrLeuTyrGlnThrLeuLeu 419
 DB 1201 GCTACATTCTACCGCTGAGTGTCTTCAAGAGGAGATATCTTATCAACCTTGGCT 1257
 RESULT 8
 AX680136 1824 bp DNA linear PAT 29-MAR-2003
 LOCUS AX680136
 DEFINITION Sequence 21 from Patent WO0181555.
 ACCESSION AX680136
 VERSION AX680136.1 GI:29369916
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Gururajan, R., Paterson, C., Recipon, S.A., Tribouley, C.M.,
 Hafalia, A., Khan, F.Y., Yue, H., Au-Young, J., Bandman, O.,
 Baughn, M.R., Borowsky, M.L., Burford, N., Burrill, J.D., Elliot, V.S.,
 Gandhi, A.R., Kearney, L., Lal, P., Lu, D.A., Lu, Y., Tang, Y.T.,
 Zingler, K.A., Griffin, J.A., Hillman, D.J., Marcus, G.A., Nguyen, D.B.,
 Policky, J., Ramkumar, J., Thornton, M., Walla, N.K. and Walsh, R.T.,
 Human kinases
 Patent: WO 0181555-A 21 01-NOV-2001;
 Incyte Genomics, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..1824
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 3107952CB1"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,056-190 Length: 1824
 Score: 2191.00 Matches: 417
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.32% Indels: 0
 DB: 6
 US-10-798-532-2 (1-419) x AX680136 (1-1824)
 QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20

```

Db      169 ATGGCCGGATCCCGCTGCGCGCGCGAGCCGCGCGCTTTCTGGAGCGCTTCGGG 228
Qy      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      229 CGGCTGGAGGATACAGACCGCTGGAGTACGGCTCTCCGCTCGGAGTACGGGTT 288
Qy      41 ArgCysArgGlyAanProGlySerProGlyValAlaLeuGlyGlnPheLeuProGly 60
Db      289 CGCTGCTGGGGGACCGCTGGCTCGCCCGCGCGCTCAAGAGATTCCTTGCGCCAGGA 348
Qy      61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
Db      349 ACCACCGGGGCTCGCGCTCTGCGCGAGTATGTTTCGCAAGAGAGGGCGGCTG 408
Qy      81 GluGlnLeuGlnGlnLysArgAanIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db      409 GAAACAGTTCAGAGGATCAGAAACATCGTACCTTTGATGAGTGTTCACATCCACTT 468
Qy      101 SerProAanValProSerArgCysLeuLeuLeuGlnLeuLeuAanPylSerValSerGlu 120
Db      469 TCTCCAAATGTGCATCAGCTGCTGTTGCTTGAACCTCGTGATGTCAGTGTTCGAA 528
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysValAlaArg 140
Db      529 TTGCTCTTATATCCAGTCCAGCGGTTGTTCCATGTGATGATACAGCACTTGCCCGA 588
Qy      141 AspValLeuGlnValAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys 160
Db      589 GATGTTTGGAGGCGCTGCTGCTTTCTTCATCATGAGGCTATGTCATGCGGACCTCMAA 648
Qy      161 ProArgAanIleLeuTrpSerAlaGluAanGlyCysPheLeuLeuLeuAspPheGlyLeu 180
Db      649 CCACTGATACATATTTGGAGTGCAGAGAAATGTTTAACTCATATGACTTGTGACTT 708
Qy      181 SerPheLysGlnGlyAanGlnAanPylValTyrIleGlnThrAspGlyTyrArgAlaPro 200
Db      709 AGCTTCAAGAGAGGCAATCAGATGTAAAGTAAATATTCAGACAAACGGGATTCGGGCTCCA 768
Qy      201 GluAlaGlnLeuGlnAanCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
Db      769 GAAAGCAATTCGAAATTCCTGCGCCAGGCTGCGCTGCAGAGTACAGATGTAAC 828
Qy      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnLeuTrpSerGlyMetLys 240
Db      829 TCAGCTGTTGATCTGTGAGAGCTTAGAATCATTTTACTGGAATGTTCTCAGGAATGAAA 888
Qy      241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAanSerSerAlaIleIleAspHis 260
Db      889 CTGAAACATACAGTCAAGATCTCAGGAATGGAAGCAAAACGTTCTCTATTTATGATCAC 948
Qy      261 IlePheAlaSerLysAlaValAlaAanAlaAlaIleProAlaTyrHisIleLeuArgAspLeu 280
Db      949 ATATTGTCAGTAAACAGTGGATGCGCGCAATTCAGCTTACCTACAGACACTT 1008
Qy      281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnLeuMetAlaLeuCys 300
Db      1009 ATCAAAAGCATGCTTCATGATGATCCAAAGAGAAATTCCTGCTGAATGCGCATGTGCG 1068
Qy      301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db      1069 AGCCCATTTCTTAAAGATTCCTTTTGGCCCTCATATGGAATGTGATGCTTCCCACT 1128
Qy      321 ProValLeuArgLeuLeuAanValLeuAspAspAspTyrLeuGlnGlnGlnGlnTyr 340
Db      1129 CCGATGCTAAGATCTGCTGAATGCTGCGATGATGATTAATTTAGAAATGAAAGGAAATAT 1188
Qy      341 GluAspValValGlnAspValLysGlnGlnCysGlnLysTyrGlyProValValSerLeu 360
Db      1189 GAAAGATGTTAGAAAGATGTAAAGAGAGAGTCAAAAATATGCAAGAGGTGATCTCTA 1248
Qy      361 LeuValProLysGlnAanProGlyArgGlyGlnValPheValGluTyrAlaAanAlaGly 380

```

```

Db      1249 CTTGTTCCAAAGGAAATCTCGGACAGAGACAAGTCTTTGTTAGTANGCAAAATCTGCT 1308
Qy      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db      1309 GATTCCAAAGCTGCGGAGAAATTTACTGACTGGAAGATGTTTGAATGGAAATTTGTGTG 1368
Qy      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db      1369 GCTACATTCCTACCGCTGAGTGCCTTACAAAGGAGATATCTGTATCAAACTTGCTT 1425

RESULT 9
RANKISRNA
LOCUS      RANKISRNA               1633 bp      mRNA      linear      ROD 08-JUL-2002
DEFINITION R.norvegicus mRNA for KIS protein.
ACCESSION  X98374
VERSION    X98374.1 GI:1403531
KEYWORDS   kis gene.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

REFERENCE
  1
AUTHORS    Maucuer,A., Ozon,S., Manceau,V., Gavet,O., Lawler,S., Curmi,P. and
            Sobel,A.
TITLE      KIS is a protein kinase with an RNA recognition motif
JOURNAL    J. Biol. Chem. 272 (37), 23151-23156 (1997)
MEDLINE    97435279
PUBMED     9287318
REFERENCE   2 (bases 1 to 1633)
AUTHORS    Maucuer,A.
TITLE      Direct Substitution
SUBMITTED  (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer
            Moulin, 75005 Paris, FRANCE
FEATURES
  source     location/Qualifiers
            1..1633
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            /cisue_type="brain"
            /clone_lib="7"
            /dev_stage="neonate"
            1..1633
            /gene="KIS"
            /gene="KIS"
            /gene="KIS"
            /gene="KIS"
            /codon_start=1
            /protein_id="CAA67021.1"
            /db_xref="GI:1403532"
            /db_xref="GOA:063285"
            /db_xref="Swiss-Prot:O63285"
            /trna_action="MAGSGCAGGAPPRFLEAFGRLMQVQSRUGSGSSASVYRVACCG
            TPSPGALKOFLPPGTGGAASAAEYGRKRRLAELOGRNVTLYGVPTTIFSP
            NPSRCILLETIDVSELYSSHQSGSMKTIQHCARDVLEALFLHHEGVHADLK
            PNIIIMSABNECFKIIDPGLSFKENQDVYKIQTHGAPRELEONCLAOGLSDTE
            CTSADVLSLGIILLEMFSKMLKRTVSSQEKANSSAIIDHI FASKVAVNAIATYH
            LKDLKSMHDEDPRIIPAEVALCSPPFSIPAPRIEDLVMLPTVRLNLVLDLYL
            ENEDYEDVEVDEKCECKYGVVSLVLPKENPGRGVFEVYANAGDSKAAQKLTGR
            MFDGRFVATFPLPASYKRGYLYQTL"

ORIGIN
Alignment Scores:
Pred. No.:      5,05e-190      Length:      1633
Score:          2186.00      Matches:      415
Percent Similarity: 99.52%      Conservative: 2
Best Local Similarity: 99.05%      Mismatches:  2
Query Match:    99.09%      Indels:      0
DB:             10      Gaps:      0

US-10-798-532-2 (1-419) x RANKISRNA (1-1633)
Qy      1 MetAlaGlySerGlyCysAlaAlaTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20

```

```

Db      39  ATGCCCGGCTCCGGCTCGCGGCGGCGGCGGCGCGGCTTCTTGAGAGCGCTTCCGGG  98
Qy      21  ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal  40
Db      99  CGCGCTGTGGCAGGTACAGAGCCGCGCTAGGACGCGGCTCCGCGCTCGGTGTACCGGGTG  158
Qy      41  ArgGlySerGlyAspProGlySerProProGlyValAlaLeuGlnGlnPheLeuProProGly  60
Db      159  CGCTGTGGCGACCTCAGAGCTCGCCCCCGGCGGCGCTCAAGAGATTCCTGCTCCGGGA  218
Qy      61  ThrThrGlyAlaAlaAlaSerAlaAlaGlyTyrGlyPheArgGlyGlnArgAlaAlaLeu  80
Db      219  ACCACCGGGGCTCGCGCTCGGTGGCGAGTATGTTTCCAAAAGAGGGAGCGCTG  278
Qy      81  GlnGlnLeuGlnGlyYHIsArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe  100
Db      279  GAGCAGTTTCAGGGTCACAGAAACATCGATCTTATACGAGTCTTTACATACACCTTC  338
Qy      101  SerProAsnValProSerArgGlyLeuLeuLeuGlnLeuLeuAspValSerValSerGlu  120
Db      339  TCTCCAAATGTGCATACGCTGTCTGTGTGAATCCTGAGTGCAGCGTTTGGAA  398
Qy      121  LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg  140
Db      399  TTGCTTGTATATTCCAGTACACAGGTTGCTCATGTGATGATCCAGCACTGTGCAGA  458
Qy      141  AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys  160
Db      459  GATGTCTGTGAGGCGCTTGTCTTTCTTTCACATGAGGCGTATGTCATGACAGCTCAAG  518
Qy      161  ProArgAsnIleLeuTrpSerAlaGlnGlnGluCysPheLysLeuIleAspPheGlyLeu  180
Db      519  CCACGAAACATCTGTGAGTGGCGAGAACAGATGCTTAAAGCTTATGACTTGGACTC  578
Qy      181  SerPheLysGlnGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro  200
Db      579  AGCTTCAAAAGAGGCAATCAGAGCGTGAATATTCAGACAGACGGGTATCGAGCTTCT  638
Qy      201  GlnAlaGlnLeuGlnGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr  220
Db      639  GAAGCAGAACTGCAGAAATGCTTGGCCGCGGCGGCTGACAGATACAGAGTATACC  698
Qy      221  SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys  240
Db      699  TCAAGCTGTGATCTCGAGCTTGGAAATCATTTTACTGAAATGTTCTCAGAAAGAA  758
Qy      241  LeuLysHisThrValArgSerGlnGluTyrPheAlaAsnSerSerAlaIleIleAspHis  260
Db      759  CTGAAACATACAGTCAAGATTCAGAGTGAAGGAAACAGTTCTGCTATATTATTCAT  818
Qy      261  IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu  280
Db      819  ATATTGGCCAGTAAACAGTGTGAATGCCGCAATTCACGCTATCACCACAGACCTT  878
Qy      281  IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuCys  300
Db      879  ATCAAAAGCATGCTTCATGACGACCCAGAGAGAGATCCCTCTGAGATGCGCTGTGC  938
Qy      301  SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr  320
Db      939  AGGCCATTTCTTTCAGCTCTTTTGGCCCTCATATTGAAGATCTGGTGAATGCTTCAACT  998
Qy      321  ProValLeuArgLeuLeuAsnValIleAspAspAspTyrLeuGlnGlnGlnGlnGlyTyr  340
Db      999  CCAAGTCTCAGACTCTCTCAATGTGCTGAGATGATGATATCTTGAATAATGAAATGATAT  1058
Qy      341  GlnAspValValGlnAspValLysGlnGlnCysGlnLysTyrGlyProValValSerLeu  360
Db      1059  GAAGAGTGTGTAAAGATGTAAAGAGAGTGTCAAGAAATATGACACAGGTGTTTCTCTG  1118
Qy      361  LeuValProLysGlnAsnProGlyArgGlyGlnValPheValGlnTyrAlaAsnAlaGly  380
Db      1119  CTTGTTCCAAAGAAATCTCTGCGACAGAGCAAGTCTTTCTTGAATACGCAACGCTGGA  1178

```

```

Qy      381  AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal  400
Db      1179  GATTCCAAAGCTGTCTCAGAAATGCTGACTGGAGAGATGTTTGAAGGAGATTTGTGTG  1238
Qy      401  AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu  419
Db      1239  GCTACATTTTACCCTGATGCTCTACAGAGGAGATATCTTTATCAAACTTGCTT  1295

RESULT 10
RNU70372      3244 bp      mRNA      linear      ROD 08-SEP-1999
LOCUS      Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
DEFINITION      complete cds.
ACCESSION      U70372
VERSION      U70372.2 GI:5821767
KEYWORDS      Rattus norvegicus (Norway rat)
SOURCE      Rattus norvegicus
ORGANISM      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
      Rattus.
REFERENCE      1 (bases 1 to 3244)
AUTHORS      Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Maine,R.E.
      and Eipper,B.A.
TITLE      Novel proteins that interact with the COOH-terminal cytosolic
      routing determinants of an integral membrane peptide-processing
      enzyme
JOURNAL      J. Biol. Chem. 271 (45), 28636-28640 (1996)
MEDLINE      97067094
PUBMED      8910496
REFERENCE      2 (bases 1 to 3244)
AUTHORS      Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
      Eipper,B.A. and Maine,R.E.
TITLE      The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
      Determinants of the Peptide Processing Enzyme Peptidylglycine
      alpha-Amidating Monooxygenase
JOURNAL      J. Biol. Chem. 274 (1999) In press
MEDLINE
PUBMED
REFERENCE      3 (bases 1 to 3244)
AUTHORS      Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Maine,R.E.
      and Eipper,B.A.
TITLE      Direct Substitution
JOURNAL      Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
      School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
      21205, USA
REFERENCE      4 (bases 1 to 3244)
AUTHORS      Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
      Eipper,B.A. and Maine,R.E.
TITLE      Direct Substitution
JOURNAL      Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
      School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
      21205, USA
REMARK      Sequence update by submitter
COMMENT      On Sep 3, 1999 this sequence version replaced gi:1698778.
FEATURES
source
1..3244
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
162..1421
/function="serine protein kinase"
/notes="P-CIP2"
/codon_start=1
/product="PAM COOH-terminal interactor protein 2"
/protein_id="AAC53031.2"
/db_xref="GI:5821768"
/translation="MAGSGCAGAPPRFLAFGLMVOYSLGSSGSAVYRVCCG
TPGSPGALGKPLPCTGALSAAEYGRKERALLPOLQGRNVTLYGVTHPS
NTPSRCLLELDVSVSELYTSSHQGSMKTIQCARVLEALAFIHHEGTVYADLK
PNNILWSAENECPKLIDFGLSPKGNQDVKYIQDGYRPALELDNCLAQAGDSGE
CTSAVDLSEIGLILEMFGMKLKTATYQEWKANSALIIDIFASKAVVNAIPAYH
LIDLKSMHDDPSRRIPALMLCSFPFSIPAPAIIDLVMPTVLRLNVLDDYL
ENDEVEDVEDVEKBCQKYGIVVSLVPEKPNRGQYVEVYANAGDSKAAQKLTGR

```

ORIGIN
MEDKRFVATYFPLSAVKRGYLYQTL"

Alignment Scores:

| Pred. No.: | 1,31e-189 | Length: | 3244 |
|------------------------|-----------|---------------|------|
| Score: | 2186.00 | Matches: | 415 |
| Percent Similarity: | 99.52% | Conservative: | 2 |
| Best Local Similarity: | 99.05% | Mismatches: | 2 |
| Query Match: | 99.09% | Indels: | 0 |
| DB: | 10 | Gaps: | 0 |

US-10-798-532-2 (1-419) x RNUT0372 (1-3244)

QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProAlaPheLeuGluAlaPheGly 20
 DB 162 ATGGCCGGGCTCCGGCTGGCGGGGCGCGAGCCGCGCGCTTCTCGAGAGGCTTCGGG 221

QY 21 ArgLeuTrpGlnValGlnSerAlaGluGlySerGlySerSerAlaSerValTyrAlaGVal 40
 DB 222 CGGCTGTGGCAGGTACAGAGCCGCTTAAGCAGCGGCTCTCGGCTCGGTGTCGGGCTG 281

QY 41 ArgCysCysGlyValAsnProGlySerProGlyValAlaLeuPheGlnPheLeuProGly 60
 DB 282 CGCTGTGGGCACTCCAGGCTGCGCCCGCGGCTCTCAAGCAGTTCTGCTCGGGA 341

QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyAlaGValAlaLeu 80
 DB 342 ACCACCGGGGCTCCGCTGCTGCGGTGCGAGTATGTTCCGAAAGAGAGGCGAGCTG 401

QY 81 GluGlnLeuGlnGlnGlnHsaArgGlnLeuValThrLeuTyrGlyValPheThrIleHsaPhe 100
 DB 402 GACGACGTGGAGGTCACAGAGAACATGTAATTAACGAGATCTTTACCATACACTTC 461

QY 101 SerProAsnValProSerArgCysLeuLeuGluLeuLeuAspValSerValSerGly 120
 DB 462 TTCGCCAATGTGCAATCAACGCTGTCTTCTTGAATCCTGATGTCAGCCTTTCGGA 521

QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 522 TTCCTGTATATTCACAGTACACAGGCTGCTCCATGTGATGATCCAGCACTGCGAGA 581

QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyrValHsaAlaAspLeuVal 160
 DB 582 GATGCTCTGGAGGCCCTTGTCTTTCTTACCATGAGGCTATATGCCATGAGACCTCAAG 641

QY 161 ProArgAsnIleLeuTyrSerAlaGluAsnGluCysPheLeuIleAspPheGlyLeu 180
 DB 642 CCACGAAACATCTCTGCGAGCTGGGAGAACGATGCTTAAAGCTTATGACTTGGACTC 701

QY 181 SerPheLeuGlnGluValangInaAspValTyrTyrIleGlnThrAspGlyTyrAlaGAlaPro 200
 DB 702 AGCTTCAAGAGAGCAATCAGAGACGTAAGTATTAATTCAGACAGCGGTATCCAGCTCCT 761

QY 201 GluAlaGluLeuGlnHsaCysLeuAlaGlnAlaGlyLeuGlnInserAspThrGluCysThr 220
 DB 762 GAAGCAGAACTGAGAAATTCCTGGCCAGCGGCTGAGAGTGTATACAGAGTGAACC 821

QY 221 SerAlaValaAspLeuTyrSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetCys 240
 DB 822 TCAGCTGTGATCTCTGGAGCCTTGGAAATCAATTTTACTGGAATGTTCTCAGAGAAAGAA 881

QY 241 LeuLeuHisThrValArgSerGlnGluTyrPylValaAsnSerSerAlaIleIleAspHis 260
 DB 882 CTGAAACATACATCAGATCTCAGAGATGAGAAAGCAAGTTCTGCTATATTAATGATCAT 941

QY 261 IlePheAlaSerIleValaValaValaAlaIleProAlaTyrHisLeuAspAlaLeu 280
 DB 942 ATATTGTCAGTAAGAGATGTGAATGCGCAATTCACCTTAATCACTCAAGAGACCTT 1001

QY 281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 1002 ATCAAAAGCATGTTTCAATGACACCCAGACAGAAAGATCCCTCTGATGATGCGCTTGTGC 1061

QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValIleLeuProThr 320
 DB 1062 AGCCCATTCCTTTAGCATTCCTTTTGGCCCTCATATGAAGATCGTGTGATCTTCCACT 1121

QY 321 ProValLeuArgLeuLeuLeuValLeuAspAspTyrLeuGlyValAsnGluGlnGluTyr 340
 DB 1122 CCAGTCTCAAGACTCTCTCAATGTGCTGGATGATGATATCTTGAATATGAAGATGAATAT 1181

QY 341 GluAspValaValaGluAspValIleGluGlnGluCysGlnIleTyrGlyProValIleSerLeu 360
 DB 1182 GAAGATGTTGTAGAAATATTAAGAAAGAGTGTCAAGAAATATGACCACTGCTTCTCG 1241

QY 361 LeuValProGlyGluAsnProGlyValGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 DB 1242 CTGTGTCAAAGGAAATCTGCGCAGAGCAAGTCTTGTGTGATGATCCCAACGCTGGA 1301

QY 381 AspSerValaAlaAlaGlnIleLeuLeuThrGlyValGlnMetPheAspGlyAspPheVal 400
 DB 1302 GATTCCAAAGCTGCTCAAGAAATTTGCTGACTGGAGAGATTTTGACGAGAAATTTGTTG 1361

QY 401 AlaThrPheTyrProLeuSerAlaTyrIleValArgIleTyrLeuTyrGlnThrLeuLeu 419
 DB 1362 GCTACATTTACCCGCTGAGTGTGCTTCAAGAGGAGTATCTTATCAAACTTGCTT 1418

RESULT 11
 BC058732
 LOCUS
 DEFINITION
 Mus musculus kinase interacting with leukemia-associated gene (setachmin), mRNA (cDNA clone MGC:64770 IMAGE:6414877), complete cds

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BC058732
 BC058732.1 GI:37194892
 MGC.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS

1 (bases 1 to 2041)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ueda, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butcherfield, Y.S., Kizylinski, M.I., Skalek, U., Small, D.E., Schermer, A., Schein, J.E., Jones, S.J., and Matre, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

2 (bases 1 to 2041)
 Strausberg, R.
 Direct Submision
 Submitted (24-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LUNL)

REMARK
 COMMENT

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu/tom-casavant@uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldi, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6754447.

FEATURES

source

1..2041
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:64770 IMAGE:641487"
 /issue_type="Brain, pool of enriched mouse brain tissues from 13.5, 14.5, 16.5 and 17.5 dpc"
 /clone_lib="NIH BMAP_FRO"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"

gene

1..2041
 /gene="Kist"
 /note="synonym: KIS"
 /db_xref="LocusID:16589"
 /db_xref="MGI:1341908"
 154..1413
 /gene="Kist"
 /codon_start=1
 /product="kinase interacting with leukemia-associated gene (statmin)"
 /protein_id="AAH58732.1"
 /db_xref="GI:37194893"
 /db_xref="MGI:1341908"
 /db_xref="LocusID:16589"

CDS

/ORIGIN
 Alignment Scores:
 Pred. No.: 8,5e-190 Length: 2041
 Score: 2185.00 Matches: 415
 Percent Similarity: 99.28% Conservative: 1
 Best Local Similarity: 99.05% Mismatches: 3
 Query Match: 99.05% Gaps: 0
 DB: 10
 US-10-798-532-2 (1-419) x BC058732 (1-2041)
 1 MetAlaG1SerG1yCySa1aTg1yAlaG1uProProArpPheLeuG1uAlaPheG1y 20
 154 ATGGCGGGGTCGGGCTGGCGGGGCGGAGCGCGGCTTCTCGAGGCGCTTCGGG 213
 21 ArgLeuTrpG1uAlaG1nSerArG1uG1ySerG1ySerSerAlaSerValTyrArG1y 40
 214 CGGCTGTGGAGCTCCAGAGCCGCTGGGCGAGCGGCTCGTGGCTCGGTATCCGGG 273
 41 ArgCyCyG1yAnpProG1ySerProProG1yAlaLeuG1nG1nG1nG1nG1nG1n 60
 274 CGGTGTGCGGTATCCCGGGCTGGCGGCTTCAAGAGTTCCTCGCTCCGGGA 333
 61 ThrThrG1yAlaAla1aSerAla1aG1uTyrG1yPheArg1yG1uArg1aAlaLeu 80

DB 334 ACCACCGGGGCTGGCGGCTCGGCGGAGGTATGTTCCGAAAGAGGGCGGCTG 393
 QY 81 GluGlnLeuGlnG1yH1aArgAn11eValThrLeuTyrG1yAlaPheThr11eH1aPhe 100
 DB 394 GAGCAGTTGACAGGTCACAGGAACCTGGACTTATATGAGGTCTTATACCTACCTTC 453
 QY 101 SerProAsnValProSerArgCySerLeuLeuG1uLeuLeuAspValSerValSerG1u 120
 DB 454 TCTCCCATGTGTCATCAGCTGTCTGCTGCTTGAATCTCGAATGTCAGTGTTCGAA 513
 QY 121 LeuLeuLeuTyrSerSerH1aGlnG1yCySerMetTrpMet11eGlnH1aCySa1aArg 140
 DB 514 TTGCTTTTATATTCAGTCATCAGGGCTGCTCAGTGGATGATACAGCAGTGGCCGA 573
 QY 141 AspValLeuG1uAlaLeuAlaPheLeuH1aG1uG1yTyrValH1aAlaAspLeu1y 160
 DB 574 GATGTTCTGGAGCGCTCCCTTCTTCTTCACATGAGGGGCTACGTCATGACAGCTCAA 633
 QY 161 ProArgAsn11eLeuTrpSerAlaG1uAnG1uCyPheLeu11eAspPheG1yLeu 180
 DB 634 CACGAAACATCTGTGAGTCGAGAGAAAGAGTGTTHAAGCTTATGACTTGGACTTC 693
 QY 181 SerPhe1yG1uG1yAsnG1nAspVal1yTyr11eGlnThrAspG1yTyrArgAlaPro 200
 DB 694 AGCTTCAAGAGAGCAATGAGAGCTGAAGATATTCAGACAGAGGGGTATGGCGTCTCT 753
 QY 201 GluAlaG1uLeuG1nAsnCySerLeuAlaG1nAlaG1yLeuG1nSerAspThrG1uCyThr 220
 DB 754 GAAGCAGAGCTCAGAACTGCTTGGCCGAGCGGCTGCAAGAGATGATCAGAAATGTAC 813
 QY 221 SerAlaValAspLeuTrpSerLeuG1y11eLeuLeuG1uMetPheSerG1yMet1y 240
 DB 814 TCAAGCTGTGATGTGGAGGCTCGGAATCATTTATCGAAATGTTCTTCAGAAATGAAG 873
 QY 241 Leu1yH1aSerThrValArgSerGlnG1uTrp1yAlaAsnSerSerAla11e11eAspH1a 260
 DB 874 CTGAAGAACATCAGTCAGATCTCAGGATGGAAGGCAACAGTTCGCTATTTATGATCAT 933
 QY 261 11ePheAlaSer1yAlaValAlaValAsnAla11eProAlaTyrH1aLeuArgAspLeu 280
 DB 934 ATATTTGCGCAGTAAGAGATGTGATGCGGAATTCAGCTTATCAGCTCAGAGCCTT 993
 QY 281 11e1ySerMetLeuH1aAspAspProSerArgAlaG1y11eProAlaG1uMetAlaLeuCy 300
 DB 994 ATCAAAAGCATCTCATGATGACCCGGGAGAGAAATCCCTGTCAGATGGCATTTGTC 1053
 QY 301 SerProPhePheSer11eProPheAlaProH1a11eG1uAspLeuValMetLeuProThr 320
 DB 1054 AGCCATTTCTTATGATCTCTTTGCGCTCATATTTGAAGATCTGTAATGCTTCGACT 1113
 QY 321 ProValLeuArgLeuLeuAnValLeuAspAspAspTyrLeuG1yAsnG1uG1uTyr 340
 DB 1114 CCAGTGCTCAGGCTCCTCATGTGCTGATGATGATTAACCTTGAATAATGAATATAT 1173
 QY 341 GluAspValValG1uAspVal1yG1uG1uCyG1n1yTyrG1yProValValSerLeu 360
 DB 1174 GAAGATGTTGAGAGATGTGAAGAGAGGTCTGAAATATGACCAATGTGTTCTCG 1233
 QY 361 LeuValPro1yG1uAnpProG1yArgG1yG1nValPheValG1uTyrAlaAsnAlaG1y 380
 DB 1234 CTGTGTCCAAAGAAATCTCTGGCAGAGACAAAGTCTTGTAAGTACCGGAAGCTGGT 1293
 QY 381 AspSer1yAlaAlaG1n1yLeuLeuThrG1yArgMetPheAspG1y1yAspValVal 400
 DB 1294 GATTCCAAAGCTGCTCAGAAATGCTGACTGGAGAGATGTTGAAGGGAAGTTGTTGG 1353
 QY 401 AlaThrPheTyrProLeuSerAlaTyr1yAsArg1yTyrLeuTyrG1nThrLeuLeu 419
 DB 1354 GCTACATTTACCGCTGAGTGTCTTACAGAGAGGAATCTTATACAACTTGTCTT 1410
 RESULT 12
 AC129141_1/c

WPCOMMENT

Sequence split into 4 fragments LOCUS AC129141 Accession AC129141

| Fragment Name | Begin | End |
|---------------|--------|--------|
| AC129141_0 | 1 | 110000 |
| AC129141_1 | 100001 | 210000 |
| AC129141_2 | 200001 | 310000 |
| AC129141_3 | 300001 | 363206 |

Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH2)

Alignment Scores:

| Pred. No.: | 1.18e-186 | Length: | 110000 |
|------------------------|-----------|---------------|--------|
| Score: | 2177.00 | Matches: | 414 |
| Percent Similarity: | 99.28% | Conservative: | 2 |
| Best Local Similarity: | 98.81% | Mismatches: | 3 |
| Query Match: | 98.69% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-10-798-532-2 (1-419) x AC129141_1 (1-110000)

QY 1 MetAlaGlySerGlyCysAlaATPGLyAlaGluProProAArgPheLeuGluAlaPheGly 20

DB 70316 ATGGCCGGGCTCCGCTGGCGGTGGGGCCGAGCCCGCGCTTCTCGAGGCTTCGGG 70257

QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40

DB 70256 CCGCTGTGGCAGGTACAGAGCCGCTTAGCGAGCGGCTCTCGGCTCGGTGTACCGGGGTG 70197

QY 41 ArgCysCysGlyValanProGlySerProProGlyValLeuValGlnPheLeuProProGly 60

DB 70196 CGCTGTGGCGGCTCCAGGCTGCGCCCGCGGCTTCAAGAGTTCGCTCGCGGGA 70137

QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80

DB 70136 ACCACGGGGCTCCGCTGGCGGTATGTTTCCGAAAGAGGGCAGGGCTG 70077

QY 81 GluGlnLeuGlnGlnYhiSarGanllleValThrLeuTyrGlyValPheThrIleHisPhe 100

DB 70076 GAGCAGTTCGAGGGGTACAGAGAACATGTCATTATACGAGCTTTCACATACACTTC 70017

QY 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120

DB 70016 TCTCCAAAGTGTGCATCACGCTGTGTTGTTGAACTCTGAGATGTCAGGTTTCGAA 69957

QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140

DB 69956 TTGCTTGTATATTCACGTACCAAGGGTGTCTCATGTGATGATACAGCATGTGCAAG 69897

QY 141 AspValLeuGlnValAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeu 160

DB 69896 GATGTTCTGGAGGCCCTTGTCTTTCTTTCACCATAGAGGCTATGTCATGACAGCTCAA 69837

QY 161 ProArgAsnIleLeuTrpSerAlaGluValanGluCysPheLeuLeuLeaPheGlyLeu 180

DB 69836 CACGAAACATCTCTGTGAGTGTGGAGAACAGAGTCTTTAAGCTTTTGACTTTGAGCTC 69777

QY 181 SerPheLeuGlnGlyValanGlnAspValTyrTyrIleGlnThrAspGlyTyrArgAlaPro 200

DB 69776 AGTTTCAAGAGACCATACAGACGTGAATATTCAGACACACGGGTATGAGCTCCG 69717

QY 201 GluAlaGlnLeuGlnAnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220

DB 69716 GAAGCGAATCTGCAGATTGCTTGGCCCGCGGCTGACAGATGATACAGAGTGTACC 69657

QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnLeuPheSerGlyMetCys 240

DB 69656 TCAAGTGTGTATCTCGAGCCTTGGATATTTACTGGAAATGTTCTCAGAGAAAGAA 69597

QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260

DB 69596 CTGAAACATACATCAGATCTCAGAGTGAAGCAACAGTTCTGTATATTATGATCAT 69537

QY 261 IlePheAlaSerLysAlaValValAlaAlaIleProAlaTyrHisLeuArgAspLeu 280

DB 69536 ATATTTCAGTAAGACAGTGTGAATGCCGAATTCAGCCTATCAGGAGACCTT 69477

QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

DB 69476 ATCAAAAGCATGTCTTCATACACACCCAGACAGATTCCTGTGAGATGGCCTTGTGC 69417

QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320

DB 69416 AGCCCATCTTTAGCATTCCTTTTGCCCTCATATGAAGATCTGGTGTGCTCCAACT 69357

QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGluTyr 340

DB 69356 CCAAGTCTAGACTCTCATATGTCTGATGATGATCTATCTGAATAAGATGAATAT 69297

QY 341 GluAspValValGluAspValLysGlnGluCysGlnLysTyrGlyProValValSerLeu 360

DB 69296 GAAAGTGTGTGAAGATTTAAAGAGAGTGTCAAAATATGACCACTGCTTCTCG 69237

QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

DB 69236 CTGTTCCAAAGGAAATCTCTGGCAGAGCAAGTCTTGTGAGTACGCAACCGCTGGA 69177

QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400

DB 69176 GATTCCAAAGCTGCTCAGAAATGTGACTGAGAGATGTTGATGGGAAGTTGTGTG 69117

QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419

DB 69116 GCTACATTTTACCCGCTGAGTGCCTTACAGAGGAGATCTTTATCAACCTTGCTT 69060

RESULT 13

MMPK1S 1254 bp mRNA linear ROD 08-JUL-2002

LOCUS Mus musculus partial mRNA for K1S protein kinase (K1S gene). Y10725

DEFINITION X10725.2 GI:21726712

ACCESSION K1S gene; protein kinase.

VERSION Y10725

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1254) Maucuer, A., Ozon, S., Manceau, V., Gavet, O., Lawler, S., Curmi, P. and Sobel, A.

AUTHORS K1S is a protein kinase with an RNA recognition motif

JOURNAL J. Biol. Chem. 272 (37), 23151-23156 (1997)

MEDLINE 97435279

PUBMED 9287318

REFERENCE 2

AUTHORS Maucuer, A.

TITLE Direct Submision

JOURNAL Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du Fer Moulin, Paris, 75005, FRANCE

REMARK 3 (bases 1 to 1254) revised by [4]

REFERENCE Maucuer, A.

AUTHORS Direct Submision

JOURNAL Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du Fer Moulin, Paris, 75005, FRANCE

COMMENT On Jul 10, 2002 this sequence version replaced gi:1806129.

FEATURES

source location/Qualifiers

1..1254

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="K22"

/issue_type="brain"

/clone_lib="lambda gt10"

/dev_stage="neonate"

1..1254

/gene="K1S"

11..>1254

CDS


```

/gene="Kis"
/codon_start=1
/product="Kis protein kinase"
/protein_id="CAA7174.2"
/db_xref="GI:21726713"
/db_xref="GOA:P97343"
/db_xref="Swiss-Prot:P97343"
/training="MAGSCAMGAPRFLAFLRUMOVSRSGSSASVRYRCCG
TPGSPGALKQFLPFGTGAASABYGRKRALBQJQHRNVTTLTGVTTHSP
NPSRCLLELDVSVSELLYSHQSCSMMLIQICARVLAFLHHEGVADLK
NPNIWASNECFKIDFGISFKENQDQVYIOTDGYRPAEALONCIAQOJQDTE
CSAVDMLGIIILEMFSKMLKHTFRQEWKANSASAIIDHIFASKAVNVAIYAH
LMDLIRKMLHDDPGRRIIPALMALCSPFSIIPAPHIDLVMLPTPLRLIYLVDDYL
EMDEYEDVDEYKECKCKYGVSLVYKENVGRQGVVEYANNQDSRAQKLTGR
MFDKRVATFPPLSAIKRGYL"

```

ORIGIN

Alignment Scores:

| Pred. No.: | 8, 34e-188 | Length: | 1254 |
|------------------------|------------|---------------|------|
| Score: | 2160.00 | Matches: | 410 |
| Percent Similarity: | 99.288 | Conservative: | 1 |
| Best Local Similarity: | 99.034 | Mismatches: | 3 |
| Query Match: | 97.914 | Indels: | 0 |
| DB: | 10 | Gaps: | 0 |

US-10-798-532-2 (1-419) x MMPKIS (1-1254)

```

QY 1 MetAlaGlySerGlyCysAlaArgGlyAlaGluProArgPheLeuGluAlaPheGly 20
DB 11 ATGGCGGGGTCGGCGCGCGTGGCGCGCGCGCGCGCGCTTCTGAGAGCGCTTCGGG 70
QY 21 ArgLeuThrGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB 71 CGGCTGTGGCAGGTCAGAGCGCGCTGGCGAGCGCTGCGCTGCGGTGATACCGGGTG 130
QY 41 ArgCysCysGlyValAsnProGlySerProGlyValAlaLeuGlnPheLeuProGly 60
DB 131 CGGTGTGGCGTACCCCGGGCTCGCCCGCGCGCTCAAGAGATTCCTGCTCCGGGA 190
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaLeu 80
DB 191 ACCACCGGGGCTCGCGCTCGCGCGCGAGTATGTTCCGCAAGAAGAGGGCGCGCTG 250
QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
DB 251 GAGCAGATTGCAGGGTCACAGAAACATCGACTTATATGAGACTTCTTACATACACTTC 310
QY 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
DB 311 TCTCCCAATGTCATCAGCTGTCTGCTGCTTGAACCTCTGATGTCTGATGTTGGAA 370
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTyrMetIleGlnHisCysAlaArg 140
DB 371 TTGCTTTTATATTCAGTATCATCGGGCTGCTCATGTGATATACAGCACTGTGCCGA 430
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeu 160
DB 431 GATGTTCTGAGGCTCTCGCTTTTTCACATGAGAGGCTCACTCATGACAGACCTCAAA 490
QY 161 ProArgAsnIleLeuTyrSerAlaGluAsnGluCysPheLeuLeuAspPheGlyLeu 180
DB 491 CCACGAAACATCTCTGAGCTGCGAGATGATGATTTTAAGTTATTAAGTTTGAAGCTC 550
QY 181 SerPheLeuGlnGlyAsnGluAspValIleValTyrIleGlnThrAspGlyTyrArgAlaPro 200
DB 551 AGCTTCAAGAAGAGCATCAGACGTAAGTATATATACAGACAGCGGATATCGCGCTCT 610
QY 201 GluAlaGlnLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerThrGlnCysThr 220
DB 611 GAGCAGAGCTGAGAACTGCTTGGCCAGAGCGCGCTGAGAGTATACAGATATAC 670
QY 221 SerAlaValAspLeuTyrSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetCys 240

```

```

DB 671 TCAGCTGTGATCTGTGAGACCTCGAATCATTTTACTGAAATGTCTCAGAAATGAAG 730
QY 241 LeuLysHisIleThrValArgSerGlnGluTyrPylAlaAsnSerSerAlaIleIleHis 260
DB 731 CTGAAGCATACAGTCACTGATCTCAGAAAGGAAAGGCAAGCAGTTCTGATTAATGATCAT 790
QY 261 IlePheAlaSerLeuAlaValAlaAsnAlaAlaIleProAlaTyrHisIleLeuArgAspLeu 280
DB 791 ATATTTCGATGAAAGCATGTGTGATGATCCGCAATTCAGCTTATCAGACACTT 850
QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
DB 851 ATCAAAAGCATCTCTCATGATGATACCCGGGAGAAATCCCTGCTGAGATGGCATTTGGC 910
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB 911 AGCCCATCTTTAGATTCCTTTTGCCTCCCTCATATATGAAGATCTGTATGCTTCGACT 970
QY 321 ProValLeuArgLeuLeuAsnValIleuAspAspAspTyrLeuGlyAsnGluGluTyr 340
DB 971 CAGTGTCTCAGGCTCTCTCATGTGCTGATGATGATTTACCTTGAAATGAAATGAAATAT 1030
QY 341 GluAspValValGluAspValIleGluGluGluCysGlnGlyTyrGlyProValValSerLeu 360
DB 1031 GAAAGATGTTGTGAAGATGTAAGAAAGAGGTGTCAAGAAATATGACACAGTGTCTCTG 1090
QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValIleValGluTyrAlaAsnAlaGly 380
DB 1091 CTTGTTCCAAAGGAAATCTCTGAGAGAGACAAAGCTTCGTGATACCGCAAGCTGCT 1150
QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
DB 1151 GATTCCAAAGCTCTCAGAGTGTCTGACTGAGAGAGATTTTGAACGGAAGTTTGTG 1210
QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeu 414
DB 1211 GCTACATCTTACCCGCTGAGTGTCTTACAAAGGGAGATATCTT 1252

```

RESULT 14
BC026046
LOCUS
DEFINITION
Hom sapiens kinase interacting with leukemia-associated gene (src-hom), mRNA (cDNA clone MGC:39391 IMAGE:4649261), complete cds.

ACCESSION
BC026046
VERSION
BC026046.1 GI:19684094
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 1367)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, S., Shemen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uebelin, T.B., Toshikuni, S., Carrincci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gumarane, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Halys, S.W., Villalón, D.K., Wuzny, D.M., Sodergren, E.U., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchwood, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitt, J., Myers, R.M., Butterfield, I.S., Krzyzanski, M.I., Skalske, U., Smalins, D.B., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
JOURNAL

PUBMED 12477932
 REFERENCE 2 (bases 1 to 1367)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabhu, Paraneh Seedi, JR Santos, Angeline Scherch, Ursula Skalska, Duane Smalins, Jeff Scott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://lml.lnl.gov>
 Series: IRAL Plate: 42 Row: f Column: 6
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21389438.
 Location/Qualifiers
 1. 1367
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:39391 IMAGE:4649261"
 /cissue_type="Eye, retinoblastoma"
 /clone_lib="NIH MGC_16"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1. 1367
 /gene="KIS"
 /db_xref="LocusID:127933"
 159. 1193
 /gene="KIS"
 /product="kinase interacting statmin"
 /protein_id="AAH26046.1"
 /db_xref="GI:19684095"
 /translation="MAGSCAMGAEPRFLEAFGRMLQVOSRLSGSSASVYRVRCG NGSPRGALIKOPLPGTGAASAAYGRKEAALQLOGRNITVLGVPTIHSP NVPSCRLLLELDVSYSEILLYSHOGCSMMNIOHCARVLDLNAITLHREGYVHADK PRTILWASNECEFKLIDFGISFEKGNQDVKTQTDGYRAPELQCLQAGLQSTP CTSADVLSIGIILLEMFGMKIKHTVRAQEMKANSATIDHIFASKAVNNAIPAYH LBDLIKSMLEDDPSRRIPAEMALCSPFSPFIPAPAHIEDLVMLPTVLRILNVLDDYLL ENEBEYEGC"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.77e-172 Length: 1367
 Score: 1993.00 Matches: 386
 Percent Similarity: 92.14% Conservative: 1
 Best Local Similarity: 91.90% Mismatches: 2
 Query Match: 90.34% Indels: 31
 DB: 9 Gaps: 1
 US-10-798-532-2 (1-419) x BC026046 (1-1367)
 QY 1 MetAlAGlySerGlyCysAlaATPGLYAlAGluProProArGPhetLeuGluAlaAheGly 20
 Db 159 ATGGCGGAGATCCGGCTGGCGCTGGAGCGCGGCGGCTTTCTGAGAGGCTTCGGG 218

QY 21 ArgLeuTrpGluValGlnSerArgLeuGlySerGlySerSerAlaSerValTyArgVal 40
 Db CGGCTGTGGCAGATGACAGAGCGGCTGGGGTACGGCTCTCCGGCTGGATGCGGGT 278
 QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuValGlnPheLeuProProGly 60
 Db CGCTGTGGGCAACCTGTGGCTGGCCCCGGCGGCTTCAGACATTCCTGGCCGAGGA 338
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyArgGlyPheArgGlyGluArgAlaLeu 80
 Db ACCACCGGGGCTGGCGCTCTGCCCCGAGTATGTTCCGAAAGAGGGCGGCTG 398
 QY 81 GlnGlnLeuGlnGlyHisArgAsnIleValThrLeuTyGlyValPheThrIleHisPhe 100
 Db GAAACGTTGACAGGTCACAGAAACATCGACTTGTATGATGAGTGTTTACATCCACTT 458
 QY 101 SerProAsnValProSerArgCysLeuLeuGluLeuLeuAspValSerValSerGlu 120
 Db TCTCCAAATGTCCATCAGCTGTCTGTGCTTGAACCTCGAGTGCAGTGTTCGGAA 518
 QY 121 LeuLeuLeuTySerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 Db TTGCTCTTATTCAGTACAGTACAGAGGTTGTTCCATGTGATATACAGCATTTGCCCCA 578
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyArgValHisAlaAspLeuTy 160
 Db GATGTTTGGAGCGCTTCCTTCTTCATCATGAGGCTATGTCATGCGGACCTCAA 638
 QY 161 ProArgAsnIleLeuTrpSerAlaGluLeuGlyCysPheValLeuIleAspPheIleLeu 180
 Db CCAGCTAACATATGTGTGGAGTGCAGAGATGATTTTAACTATGACTTGGACTT 698
 QY 181 SerPheLeuGluGlyAsnGlnAspValIleTyArgIleGlnThrAspGlyTyArgAlaPro 200
 Db AGCTTCAAAGAGGCAATCAGATGTAAGTATATTCAGACAGAGCGGTATGGCTCA 758
 QY 201 GlnAlaGluLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 Db GAAGCAGAAATTCAAATATGCTTGAGCCAGGCTGCGCAGAGATACAGAAATGTACC 818
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetIys 240
 Db TCAGCTGTGATCTGTGAGCTGAGCAATCATTTTACGGAATGTCTCAGAAATGAAA 878
 QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
 Db CTGAAACATACAGTACATCTCAGAAATGGAAGCAACAGTTCGTACTATTTGATCAC 938
 QY 261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyHisIleLeuArgPheLeu 280
 Db ATATTTCGACAGAAAGCAGATGTGAATCCCGCAATTCACAGCTTACACTAAGACCTT 998
 QY 281 IleLysSerMetLeuHisAspAspProSerArgGlyIleProAlaGluMetAlaLeuTy 300
 Db ATCAAAAGCATCTTCATGATGATCCAGCAAAATAATTCGTGAATGCAATTTGCTGC 1058
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db AGCCCATCTTACGATTCCTTTTCCCTCATATTTGAAGATCTGTGATGCTTCCACT 1118
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyLeuGlyAsnGluGluGluTy 340
 Db CCAGTGTCAAGACTGCTGATGTGCTGATGATGATTTCTTGAAGAAAGAGAGAAATAT 1178
 QY 341 GluAspValValGluAspValIleGluGluCysGlnTyTyGlyProValValSerLeu 360
 Db 1178 ----- 1178
 QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyArgAlaAsnAlaG 380
 Db 1179 -----GAGGTCTTTGTGATGATGCAAAATGCTGG 1208

Qy 380 yAepSerLySaLaAGlnTyrLeuSerThrglyArMetPheAspGlyLysPheVala 400
Db 1209 TGAATTCAGAGCTGCGCAGAAATTTACTGACTGAGAGATTTTGGAGGAGTTTGTGT 1268
Qy 400 lAlaThrPheTyProLeuSerAlaTyTyLysArgGlyTyLeuTyGlnThrLeuLeu 419
Db 1269 GGCTACATCTTACCCCGTGAAGCTCTACAGAGGGGATATCTGTATCAAACTTGTCTT 1326

RESULT 15
HSM805327 2940 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens mRNA, cDNA DKFZp6710511 (from clone DKFZp6710511).
DEFINITION AL834136
ACCESSION AL834136.1 GI:21739600
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2940)
REFERENCE 1 (bases 1 to 2940)
AUTHORS Koehner,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the CDNA sequencing consortium of the German
Genome Project.
This clone (DKFZp6710511) is available at the RZPD in Berlin.
Please contact the RZPD: Resourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cdna/>.
Location/Qualifiers
1..2940
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp6710511"
/db_xref="taxon:9606"
/map="11"
/clone="DKFZp6710511"
/issue_type="lymph node"
/clone_id="667 (synonym: hlnoz). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
2923

polYA_site
ORIGIN

Alignment Scores:
Pred. No.: 6 49e-163 Length: 2940
Score: 1893.50 Matches: 389
Percent Similarity: 63.46% Conservative: 0
Best Local Similarity: 63.46% Mismatches: 2
Query Match: 85.83% Indels: 223
Gaps: 1
DB: 9
US-10-798-532-2 (1-419) x HSM805327 (1-2940)

Qy 29 LeuGlySerGlySerSerAlaSerValTyTyArgValArgCySgLyAsnProGlySer 48
Db 1 CTGGGTAGAGGGCTCTCGCTCGGTGTATCGGGTTCTGCTGCGGCAACCTGTGCTCG 60
Qy 49 ProProGlyValAlaLeuLysGlnPheLeuProProGlyThrThrglyAlaAlaAlaSerAla 68
Db 61 CCCCCCGGCGCCCTCAGCAGGTTCTTGCCGCCAGGAACACCGGGGCTGCGGCTCTGCC 120
Qy 69 AlagLtyrGlyPheArgLysGlnUArgAlaAlaLeuGlnGlnGlnGlnGlnGlnGln 88
Db 121 GCGGAGTATGTTTCCGCAAGAGAGGGCGCGCTGGAACAGTTGCGAGGTCACAGAAC 180
Qy 89 lLeValThrLeuTyrglyValAlaPheThrIleHisPheSerProAsnValProSerArgCyS 108
Db 181 ATCGTACCTTGTATGAGAGTGTATTAACAATCACTTTTCTCCAAATGTCATCAGCGTGT 240

Qy 109 LeuLeuLeuGlnLeuLeuAspValSerValSerGlnLeuLeuTySerSerHisGln 128
Db 241 CTGTTGCTGAACCTCTGGATGTCAAGTGTGTGGAAATGCTCTTAATTCAGTCAACAG 300
Qy 129 GlyCySerMetTrpMetIleGlnHisCySaLaArgAspValLeuGlnAlaLeuAlaPhe 148
Db 301 GGTTGTTCATGTGATGATGATACAGCATTTGTGCGGAGATGTTTGGAGGCGCTTGTCTT 360
Qy 149 LeuHisHisGlnGlyTyTyValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAla 168
Db 361 CTTCATCATGAGAGGCTATGCTCATGCGGACCTCAACACCGTAAACATATTTGGAGTGA 420
Qy 169 GluAsnGluCyPheLysLeuIleAspPheGlyLysSerPheLysGlnGlnGlnGlnGln 187
Db 421 GAGAAATGAATGTTTAACTCATGACTTGTGACCTTGAAGGCAATCA-CGT 479
Qy 187 ----- 187
Db 480 AAGAAATAACCTTTCTTCTCTGCACTTAAATGATGCTTGAAGCCAAATATTT 539
Qy 187 ----- 187
Db 540 TTAATGACAGGTAAGTTAGTAGGGCTTCAAGTCACTTTAACTPAATTACTGAA 599
Qy 187 ----- 187
Db 600 TGCCAGGGAGTACAGTAATCAAAAAGATTGATGCCACCGTAAAAAAGACATTA 659
Qy 187 ----- 187
Db 660 TGTTTGGTCAGTAGAGAGAGCAGTAGAGATGAGCTGTGAATTTAAAGTAC 719
Qy 187 ----- 187
Db 720 TTAACCTACATGAATGTTTTTTAAATTTATNGCGTGTGGCTAATGCTGAGTTACT 779
Qy 187 ----- 187
Db 780 ATTCTGCTCTCTTCAAGTGTGGGCTGTATCTCGTGCAAGATGATGATGACTTGG 839
Qy 187 ----- 187
Db 840 CCAATACCTCTTTCCTCCATGATTTATCATTTCTTGCAAGTCTCTTATATCAGACAGT 899
Qy 187 ----- 187
Db 900 CCTAATAGTAAGAGACATCAACAAAGACTTCTTATAGCGCTATATGAAATGTAC 959
Qy 187 ----- 187
Db 960 AGCTTGCTCTGAGAACTGCCAGTTAATGAACAAAGATATGTGTAATCAAGACTTA 1019
Qy 187 ----- 187
Db 1020 AAATTTTAGCATTTAGCTTGGCAGTGGGTTTACTGCACTTTAGGGAAGTAATTCAC 1079
Qy 187 ----- 187
Db 1080 GTAGAGATGATAGAAAGGAGACAGCTTTTATTGTTGTATATTTACTCATCTTTT 1139
Qy 188 ---AspValLysTyTlleGlnThrAspGlyTyTyArgAlaProGlnAlaGlnLeuGlnAsn 206
Db 1140 TAGGATGTAAATATATTTACAGACAGGGTATCGGGCTCCAGAAAGCAAAATTTGCAAAAT 1199
Qy 207 CysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlnCyTySerAlaValAspLeuTrp 226
Db 1200 TGCTTGCCCAAGGCTGCGCTGAGAGTATACAGATGATGACTTCAGCTTTTATCTGTGG 1259
Qy 227 SerLeuGlyIleIleLeuLeuGlnUetPheSerGlyMetLysLeuLysHisThrValArg 246
Db 1260 AGCTTAGAATCATTTTACTGAGAAATGTTCTCAGGAATGAACACTGAACATACAGTACAGA 1319

```
QY 247 SerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAla 266
    |||
Db 1320 TCTCAGGAATGAGAGGAAACAGTTCGTATTGATTCACATATTGCCAGTAAAGCA 1379
    |||
QY 267 ValValAsnAlaAlaIleProAlaIleHisLeuArgAspLeuIleLysSerMetLeuHis 286
    |||
Db 1380 GTGGTAAATGCCCAATTCAGGCTTACACCTTAAGAGACCTTATCAAAAGCATGCTTCAT 1439
    |||
QY 287 AspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIle 306
    |||
Db 1440 GATGATCCAGACGAGAGAAATTCCTGCTGAAATGCGATTGTGCAAGCCCATTTAGCATT 1499
    |||
QY 307 ProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeu 326
    |||
Db 1500 CCTTTGCCCTCATATTGAAGATCTGTGATGCTTCCACTCCAGTGTAGAGACTGCTG 1559
    |||
QY 327 AsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAsp 346
    |||
Db 1560 AATGTCTGATGATGATGATTTATCTTGAGAAATGAAGAGAAATATGAGATGTTGTAGAAAGAT 1619
    |||
QY 347 ValLysGluGluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsn 366
    |||
Db 1620 GTAAAGAGAGAGTGTCAAAATATGACACAGTGTATCTTACTTGTCCAAAGGAAAT 1679
    |||
QY 367 ProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaAlaGln 386
    |||
Db 1680 CCTGGCAGAGAGACAGTCTTTGTTGATGATGCAAAATGCTGTGATTCCAAAGCTGCGCAG 1739
    |||
QY 387 LysLeuLeuThrGlyArgMetPheAspGlyLysPheValAlaIleThrPheTyrProLeu 406
    |||
Db 1740 AAATTACTGACTGAAAGATGTTGATGAGGAGTTGTTGTGGCTACATTCTACCCGCTG 1799
    |||
QY 407 SerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
    |||
Db 1800 AGTGCTACAGAGGGGATATCTGTATCAAACTTGCTT 1838
    |||
```

Search completed: November 29, 2004, 18:10:34

Job time : 4620 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 502 Seconds
(without alignments)
4381.489 Million cell updates/sec

Title: US-10-798-532-2
Perfect score: 2206
Sequence: 1 MAGSGCAMEAPRRFLIARG.....VATFPLSAKRGVLYQTLL 419

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 413486 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgm2.1/USPTO.spool.p/US10798532/runat.29112004.112452.15882/app.query.faeta_1.1166
-DB=N Geneseq.23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPT=0
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532.0CGN.1.1.886.0runat.29112004.112452.15882 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.23Sep04.*
1: geneseqn1980s.*
2: geneseqn1980s.*
3: geneseqn2000s.*
4: geneseqn2000s.*
5: geneseqn2000s.*
6: geneseqn2000s.*
7: geneseqn2000s.*
8: geneseqn2000s.*
9: geneseqn2000s.*
10: geneseqn2000s.*
11: geneseqn2000s.*
12: geneseqn2000s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 2206 | 100.0 | 1260 | AAZ51355 | AAZ51355 Human wtl |
| 2 | 2203 | 99.9 | 1260 | AAZ51356 | AAZ51356 Human tra |
| 3 | 2198 | 99.6 | 1260 | AAI68829 | AAI68829 Human pro |
| 4 | 2198 | 99.6 | 2008 | ABQ9383 | ABQ9383 Human cod |
| 5 | 2198 | 99.6 | 2622 | AAI68828 | AAI68828 Human pro |
| 6 | 2198 | 99.6 | 4065 | AAH72714 | AAH72714 Human cer |

| | | | | | |
|----|--------|------|------|----------|--------------------|
| 7 | 2198 | 99.6 | 4065 | ADL45238 | ADL45238 Human ova |
| 8 | 2191 | 99.3 | 1260 | AA506734 | AA506734 Polynucle |
| 9 | 2191 | 99.3 | 1824 | AAI18818 | AAI18818 Human kin |
| 10 | 2186 | 99.1 | 3244 | ADBS2857 | ADBS2857 Primary r |
| 11 | 2045.5 | 92.7 | 3998 | ADQ24401 | ADQ24401 Human sof |
| 12 | 789 | 35.8 | 529 | AAH70764 | AAH70764 Human cer |
| 13 | 766.5 | 34.7 | 722 | AAAO2535 | AAAO2535 Human col |
| 14 | 729 | 33.0 | 470 | AAH71966 | AAH71966 Human cer |
| 15 | 636 | 28.8 | 461 | AAH69101 | AAH69101 Human cer |
| 16 | 559.5 | 25.4 | 411 | ABQ99225 | ABQ99225 Human ORF |
| 17 | 361 | 16.4 | 231 | AAAO1259 | AAAO1259 Human col |
| 18 | 311.5 | 14.1 | 419 | ADL73740 | ADL73740 Human ova |
| 19 | 311.5 | 14.1 | 419 | ADL67345 | ADL67345 Human ova |
| 20 | 287 | 13.0 | 546 | ABO53127 | ABO53127 Oligonuc1 |
| 21 | 287 | 13.0 | 546 | ABO53126 | ABO53126 Oligonuc1 |
| 22 | 287 | 13.0 | 546 | ABQ44465 | ABQ44465 Oligonuc1 |
| 23 | 287 | 13.0 | 547 | ABQ44464 | ABQ44464 Oligonuc1 |
| 24 | 288 | 11.7 | 546 | ABO53129 | ABO53129 Oligonuc1 |
| 25 | 288 | 11.7 | 546 | ABO53128 | ABO53128 Oligonuc1 |
| 26 | 288 | 11.7 | 547 | ABO44463 | ABO44463 Oligonuc1 |
| 27 | 288 | 11.7 | 547 | ABO44462 | ABO44462 Oligonuc1 |
| 28 | 252 | 11.4 | 1065 | ADD21399 | ADD21399 Golden ha |
| 29 | 250 | 11.3 | 442 | ACH28557 | ACH28557 Human adu |
| 30 | 249.5 | 11.3 | 1400 | ADD21407 | ADD21407 Rat cycli |
| 31 | 248 | 11.2 | 1708 | ADD21395 | ADD21395 Mouse cyc |
| 32 | 243.5 | 11.0 | 2472 | AA455716 | AA455716 Arabidops |
| 33 | 243 | 11.0 | 1297 | ADD21393 | ADD21393 Human cyc |
| 34 | 243 | 11.0 | 2213 | ABR84073 | ABR84073 Human cDN |
| 35 | 243 | 11.0 | 2213 | ADD21403 | ADD21403 Chinese h |
| 36 | 243 | 11.0 | 2213 | ADL31711 | ADL31711 Human cDN |
| 37 | 243 | 11.0 | 2213 | ADH17138 | ADH17138 Human cdc |
| 38 | 243 | 11.0 | 2328 | ACC72837 | ACC72837 Human can |
| 39 | 243 | 11.0 | 2734 | ADQ25034 | ADQ25034 Human sof |
| 40 | 240 | 10.9 | 1209 | ADD21401 | ADD21401 Golden ha |
| 41 | 239.5 | 10.9 | 1335 | AAV71073 | AAV71073 CDK2-gree |
| 42 | 239 | 10.8 | 1297 | ABO60787 | ABO60787 Human 10C |
| 43 | 239 | 10.8 | 1297 | ADD21391 | ADD21391 Human cyc |
| 44 | 238.5 | 10.8 | 1650 | ADL45326 | ADL45326 Rice isop |
| 45 | 237.5 | 10.8 | 1344 | ADD21409 | ADD21409 Rat cycli |

ALIGNMENTS

| | | |
|----------|--|---|
| RESULT 1 | AAZ51355 | AAZ51355 standard; DNA, 1260 BP. |
| ID | AAZ51355 | |
| XX | AAZ51355; | |
| AC | | |
| XX | | |
| DT | 06-JUN-2000 | (first entry) |
| XX | | |
| DE | Human wild type serine/threonine kinase KIS (hKIS) gene. | |
| XX | | |
| KW | KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase; | |
| KW | dominant gene; cyclin-dependent kinase inhibitor; CKI; p27; modulator; | |
| KW | treatment; cell proliferative disease; vascular disorder; gene therapy; | |
| KW | atherosclerosis; restenosis; ds. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| FH | | |
| FT | Key | Location/Qualifiers |
| FT | CDS | 1..1260 |
| FT | | /tag= a |
| FT | | /product= "Human wild type KIS protein" |
| FT | | /transl_except= (pos:556..561, Asn) |
| XX | | |
| PN | MO20001165-Al. | |
| XX | | |
| PD | 02-MAR-2000. | |
| XX | | |
| PF | 20-AUG-1999; | 99WO-US018903. |
| XX | | |
| PR | 21-AUG-1996; | 98US-0097710P. |

XX (NABEL/) NABEL G J.
 PA (NABEL/) NABEL E G.
 XX Nabel GJ, Nabel EG;
 PI Nabel GJ, Nabel EG;
 XX MPI; 2000-237648/20.
 DR P-PSDB; AAY70305.
 XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
 PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 PT proliferation.
 XX
 PS Claim 1; Page 56; 709p; English.
 CC The present DNA sequence encodes a wild type human KIS (hKIS), isolated
 CC from a human B-cell library, using a yeast two hybrid screening system.
 CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
 CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
 CC that controls cell proliferation and is localised predominantly in the
 CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
 CC G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
 CC The hKIS sequences are used to modulate cell proliferation and treat cell
 CC proliferative and vascular diseases. The polynucleotide sequence may be
 CC used in gene therapy to treat vascular disorders such as restenosis or
 CC atherosclerosis
 CC
 SQ Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.75e-209 Length: 1260
 Score: 2206.00 Matches: 419
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-10-798-532-2 (1-419) x AAZ51355 (1-1260)

QY 1 MetAlG1SerG1yCysAlaTTPG1yAlaGluProProArpPheLeuGluAlaPheGly 20
 Db 1 ATGGCGGAGATCCCGCTGCGCTGGGGGCGGAGCCGCGCTTTCTGGAGGCTTCGGG 60
 QY 21 ArgLeuTTPG1yAlaGlnSerArgLeuG1ySerG1ySerSerAlaSerValTYrArGVal 40
 Db 61 CGGCTGAGGAGGTACAGAGCCGCTGGGTAGCGGCTCCCGCTCGGTATCGGGTT 120
 QY 41 ArgCysCysG1yAsnProG1ySerProProG1yAlaLeuLysGlnPheLeuProProG1y 60
 Db 121 CGCTGCTGGCGCAACCTGGCTGCGCCCGCGGCGCTTCAGACAGTCTTGGCCAGGA 180
 QY 61 ThrThG1yAlaAlaAlaSerAlaAlaGluTYrG1yPheArpLysGluArGAlaAlaLeu 80
 Db 181 ACCACCGGGGCTCGGCTCTGCGCGCGAGTATGTTCCGCAAGAGAGGCGGCGCTG 240
 QY 81 GluGlnLeuGlnLysIstArgAsnIleValThrLeuTYrG1yValPheThrIleAspHe 100
 Db 241 GAACAGTTGAGGGTACAGAAACATCGTGACTTTGTATGAGAGTGTTCACATCCACTTT 300
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
 Db 301 TCTCCAAATGTGCATCACCGCTGTGTGCTTGAACCTCTGATGTCAAGTGTTCGGA 360
 QY 121 LeuLeuLeuTYrSerSerIstGlnG1yCysSerMetTTPMetIleGlnIstCysAlaArg 140
 Db 361 TTGCTCTTATATTCAGTCAACAGGGTGTTCATGTGATGATGATGATGATGATGATG 420
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisIstGluG1yTYrValHisAlaAspLeuLys 160
 Db 421 GATGTTTGGAGGCGCTGTGCTTTCTTCATCAAGAGGGGATATCTCATGCGACCTCAA 480
 QY 161 ProArgAsnIleLeuTTPSerAlaGluArgGlnCysPheLysLeuIleAspPheGlyLeu 180

Db 481 CCACGTAAATATGTGAGTGCAGAGATGATGATGATGATGATGATGATGATGACTT 540
 QY 181 SerPheLysGluG1yAsnGlnAspValLysTYrIleGlnThrAspG1yTYrArGAlaPro 200
 Db 541 AGCTTCAGAGAGGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 600
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaG1yLeuGlnSerAspThrG1yCysThr 220
 Db 601 GAAGCAGATTCAGAAATGCTTGGCCGAGCTGCGCTGCAGAGATGATGATGATGATG 660
 QY 221 SerAlaValAspLeuTTPSerLeuG1yIleIleLeuLeuGluMetPheSerG1yMetLys 240
 Db 661 TCAGCTGTGATCTGTGAGCTGAGATGATGATGATGATGATGATGATGATGATGATG 720
 QY 241 LeuLysHisIstThrValArgSerGlnGluTYrLysAlaAsnSerSerAlaIleIleAspHis 260
 Db 721 CTGAAACATACAGTCAAGATCTCAGAAATGAGAGGCAACAGTTCCTGCTATTTATGATC 780
 QY 261 IlePheAlaSerLysAlaValAlaAlaAlaIleProAlaTYrHisIleLeuArgAspLeu 280
 Db 781 ATATTGGCAGTAAAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 840
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 Db 841 ATCAAAAGCATCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 901 AGCCCATCTTTAGCATTCCTTTTCCCTCATATTTGAAGATCTGTGATGCTTCCACT 960
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTYrLeuG1yAsnGluGluTYr 340
 Db 961 CCAGTGCTAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 QY 341 GluAspValValGluAspValLysGluG1yCysGlnLysTYrG1yProValLysSerLeu 360
 Db 1021 GAAGATGTTGAGAGATGTAAGAGAGAGTCAAAATATGACACAGTGTATCTCTTA 1080
 QY 361 LeuValProLysGluAsnProG1yArgG1yGlnValPheValGluTYrAlaAsnAlaGly 380
 Db 1081 CTGTTTCCAAAGAAATCTTCGACAGAGACAGCTTTGTGATATGCAATCTGGT 1140
 QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrG1yArgMetPheAspG1yLysPheValVal 400
 Db 1141 GATTCCAAAGCTGCCAGAAATTAATGACTGAGAGAGATGTTGATGAGGAAAGTTTGTG 1200
 QY 401 AlaThrPheTYrProLeuSerAlaTYrLysArgG1yTYrLeuTYrGlnThrLeuLeu 419
 Db 1201 GCTACATTTACCCGCTGAGTCTTAAGAGAGGAGATATCTGATCAAACTTGTCTT 1257

RESULT 2
 AAZ51356
 ID AAZ51356 standard; DNA; 1260 BP.
 AC AAZ51356;
 XX
 DT 06-JUN-2000 (fixet entry)
 DT
 XX
 DE Human transdominant mutant serine/threonine kinase KIS (hKIS) gene.
 XX
 KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
 KW transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27;
 KW modulator; treatment; cell proliferation disease; vascular disorder;
 KW gene therapy; atherosclerosis; restenosis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /tag= a
 FT /product= "Human transdominant mutant KIS protein"
 FT /transl_except= (pos:556..561, Asn)
 FT misc_feature 161

```

FT      /+tag= b
FT      /note= "A to G substitution resulting in a Lys to Arg
FT      change in the amino acid sequence"
XX
XX      WO200011165-A1.
XX
XX      02-MAR-2000.
XX
XX      20-AUG-1999; 99MO-US018903.
XX
XX      21-AUG-1998; 98US-0097710P.
XX
XX      (NABE/) NABEL G J.
XX      (NABE/) NABEL E G.
XX
XX      Nabel GJ, Nabel EG;
XX
XX      MPI; 2000-237648/20.
XX      P-PSDB; AAY70306.
XX
XX      Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
XX      for inhibiting the cyclin kinase inhibitor p27, and so alter cell
XX      proliferation.
XX
XX      Claim 7; Page 59; 70pp; English.
XX
XX      The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
XX      constructed by site directed mutagenesis. A single nucleotide
XX      substitution (A to G) results in a lysine to arginine change in the
XX      protein sequence. hKIS is a serine/threonine kinase, that acts as an
XX      inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
XX      controls cell proliferation and is localized predominantly in the
XX      nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
XX      G1 phase. The hKIS sequences are used to modulate cell proliferation and
XX      treat cell proliferative and vascular diseases. The polynucleotide
XX      sequence may be used in gene therapy to treat vascular disorders such as
XX      restenosis or atherosclerosis
XX
XX      Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 9,44e-209 Length: 1260
XX      Score: 2203.00 Matches: 418
XX      Percent Similarity: 100.00% Conservative: 1
XX      Best Local Similarity: 99.76% Mismatches: 0
XX      Query Match: 99.86% Indels: 0
XX      DB: 3 Gaps: 0
XX
XX      US-10-798-532-2 (1-419) x AA251356 (1-1260)
XX
QY      1 MetAlaGlySerGlyCybAlaATrPGlyAlaGluProProArpPheLeuGluAlaPheGly 20
DB      1 ATGGCGGGAATCCGCGCTGCGCTGGGGGCGCGAGCGCGCGCTTTCGAGGCGCTTCGCGG 60
QY      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
DB      61 CGGCTGTGGAGGTACAGACCGGCTGGAGGAGGAGCTTCCTCGCTCGGTGATCGGGT 120
QY      41 ArgCySeGlyAsnProGlySerProGlyAlaLeuLysGlnPheLeuProProGly 60
DB      121 CGCTGGTGGGGCAACCTGGCTGGCCCGCGGCGCTTCAGGGAGTCTTGCGCGCAGGA 180
QY      61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluAlaAlaLeu 80
DB      181 ACCACCGGGGCTCGGCTCTGCGCGCGAGTATGTTTCCGCAAAAGAGAGGGCGCTG 240
QY      81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
DB      241 GAACAGTTGCAAGGCTCACAAACATGTAATTGATGAGAGTGTTCATCAATCCACTTT 300
QY      101 SerProAsnValProSerArgCybLeuLeuGlnLeuLeuAspValSerValSerGlu 120
DB      301 TTCCAATGTGCATACGCTGTCTGTCTTGAACCTCCGATGTCAGTGTTCGGA 360

```

```

QY      121 LeuLeuLeuTyrSerSerHisGlnGlyCySeSerMetTrpMetIleGlnHisCybAlaArg 140
DB      361 TTGCTCTTATATTCACGTACACAGGTTTTCATGTGATGATACAGATTCGCCCA 420
QY      141 AspValLeuGluValAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys 160
DB      421 GATGTTTGGAGGCCCTTCTTTCTTCATCATAGAGGCTATGTCCATCGGACCTCAAA 480
QY      161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCybPheLysLeuIleAspPheGlyLeu 180
DB      481 CCACGTACATATTTGGAGTGCAGAGATGATGATTTAACTCATTAACCTTGGACTT 540
QY      181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
DB      541 AGCTTAAAGAGAGGCAATCAGATGTAAAGTATTTACAGACAGCGGATTCGGCTCA 600
QY      201 GluAlaGluLeuGlnAsnCybLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
DB      601 GAAGCAGAATTCGAAATTTGCTGGCCAGGCTGGCTTCAGAGTGAACAGAAATGTACC 660
QY      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
DB      661 TCAGCTGTGATCTGTGGAGCCTAGGAATCATTTTACTGAAATGTTTCAGGAATGAAA 720
QY      241 LeuLysHisThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHis 260
DB      721 CTGAACCATACAGTCAATCTCGATCTCGAATGAAAGCAACAGTTCCTTATTATATCAC 780
QY      261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisIleLeuArgAspLeu 280
DB      781 ATATTTCAGATTAACATGATGATGATGCGGCAATTCAGCCATATCCTTAAGAGACTT 840
QY      281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
DB      841 ATCAAAAGCATGCTTCATATATATATCCAAACAGAAATCTGCTAAATGCGATTGTGC 900
QY      301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB      901 AGCCCATTCCTTAAGCATCTCTTGTGCGCTCATATTTGAAGATCTGTATCTCCACT 960
QY      321 ProValLeuArgLeuLeuAsnValLeuAspAspTyrLeuGlyAsnGluGlnGlyTyr 340
DB      961 CCAAGTCTTAAGACTGCTGATGATGCTGATGATGATTAATCTTGGGAATGAAGCAATAT 1020
QY      341 GluAspValValGluAspValLysGlnGlyCysGlnLysTyrGlyProValValSerLeu 360
DB      1021 GAAGATGTTGTAGAAATGTAAAGAGAGTGTCAAAAATATAGACCACTGTATCTCTA 1080
QY      361 LeuValProLysGlnAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
DB      1081 CTGTGTTCAAGAGAAATCTTGCAAGAGCAAGATTTTGTGAGTATCAAAATGCTGGT 1140
QY      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
DB      1141 GATTCCAAAGCTCGGAGAAATTAATCACTGGAAGATTTTGAATGCGAAGTTGTGTG 1200
QY      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
DB      1201 GCTACATTTCTACCGCTGAGTGTCAAGAGGGGATATCTGATCAAAACTTGTCTT 1257

```

RESULT 3
 AA166829
 ID AA166829 standard; cDNA, 1260 BP.
 XX
 AC AA166829;
 XX
 XX 07-JAN-2002 (first entry)
 DE Human protein kinase polypeptide 13887 coding sequence.
 XX Protein kinase, 3714; 16742; 23546; 13887; cancer; bone disorder; human;
 KM cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatocytrophic;
 XX

KW gene therapy; ss.
 XX Homo sapiens.
 XX
 XX
 FT CDS Location/Qualifiers
 FT 1..1260
 FT /-tag= a
 FT /product= "protein kinase 13887"
 XX
 XX MO200173050-A2.
 XX
 XX 04-OCT-2001.
 XX
 XX 23-MAR-2001; 2001WO-US009483.
 XX
 XX 24-MAR-2000; 2000US-0191846P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Meyers R;
 XX
 XX WPI; 2001-61632/70.
 XX P-PSDB; AAG65767.
 DR
 XX
 XX
 PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
 PT useful in diagnosis of cancer or cellular proliferation or
 PT differential disorders and to screen for polypeptide modulators useful
 PT to treat such conditions.
 XX
 XX
 RS Claim 1; Fig 10A-C; 169pp; English.
 XX
 XX The invention provides novel human protein kinase polypeptides, 3714,
 CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
 CC protein kinase polypeptides can be expressed by standard recombinant
 CC methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
 CC are useful for diagnostic and screening methods to identify subjects (at
 CC risk of) having cancer or cellular proliferation and/or differentiation
 CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
 CC modulators are useful for the treatment of cancer, particularly colon
 CC cancer or cellular proliferation and/or differentiation disorders. Other
 CC disorders associated with 3714, 16742, 23546 or 13887 expression or
 CC activity that can be treated include bone related disorders, inflammatory
 CC disorders, autoimmune diseases, cardiovascular disorders and liver
 CC diseases. The present sequence represents a human protein kinase
 CC polypeptide 13887 coding sequence
 CC
 SQ Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2..96e-208 Length: 1260
 Score: 2198.00 Matches: 418
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.64% Indels: 0
 DB: 4 Gaps: 0
 US-10-798-532-2 (1-419) x AA166829 (1-1260)
 QY 1 MetAlaG1SerG1CySaAlaTPrG1YAlaG1uProProArpPheUeuG1uAlaPheG1Y 20
 Db 1 ATGGCCGGATCCGGCTCGCGCTGGGGCGCGAGCCCGCGCTTTCTGAGGCGCTTCGGG 60
 QY 21 ArgLeuTPrG1nValG1nSerArGLeuG1SerG1YSerSerAlaSerValTYrArGVal 40
 Db 61 CGGCTGTGGCAGGTACGAGCCGCTGTGGTACGGCGCTCCGCTCGGTGTATCGGGTT 120
 QY 41 ArgCyCyG1uYAsnPrG1YSerPrG1YAlaLeuYAsnInPheLeuPrPrG1Y 60
 Db 121 CGCTGCTGGCGCAACCTGTGCTGCCCGCGGCGCTCAAGAGATTCTTGGCGCAGGA 180
 QY 61 ThrThG1YAlaAlaAlaSerAlaAlaG1uTYrG1YPhaArgYAsG1uArGAlaAlaLeu 80
 Db 181 ACCACCGGGGCTCGCGCTCTGCGCGCGAGTATGTTTCCGCAAGAGAGGGCGGCGCTG 240

QY 81 GluG1nLeuG1nG1YnH1sArgAn11eValThrLeuTYrG1YValPheThr11eH1sPhe 100
 Db 241 GAACAGTTCCAGGTCACAGAAACATCGACCTTGTATGAGGTGTTTCAATCCACTTT 300
 QY 101 SerProAnValProSerArgCySLeuLeuG1uLeuLeuAspValSerValSerG1u 120
 Db 301 TCTCCAAATGTCACACGCTGTCTGTGCTTGAATCTCTGAATGTCAGTGTTCGGGA 360
 QY 121 LeuLeuLeuTYrSerSerH1sG1nG1YCySerMetTrpMet11eG1nH1sCySaAlaArg 140
 Db 361 TTGCTCTTATATTCACAGTACAGAGGTTGTCTCATGTGATATACAGATGTTGCCGA 420
 QY 141 AspVal1LeuG1uAlaLeuAlaPheLeuH1sH1sG1uG1YTYrValH1sAlaAspLeuYs 160
 Db 421 GATGTTTGGAGGCCCTTCTTTCTTCATCATGAGGCGCTATGTCATCGGACCTCAAA 480
 QY 161 ProArgAn11eLeuTrpSerAlaG1uAnG1uCyPheLeuLeu11eAspPheG1YLeu 180
 Db 481 CCACGTACATATGTGGAGTGCAGAGAAATGAATGTTTAACTCATTTGACTT 540
 QY 181 SerPheYAsG1uG1YAsnG1nAspVal1YsTYr11eG1nThAspG1YTYrArGAlaPro 200
 Db 541 AGCTTCAAAGAAAGCAATCAGATGTAAAGTATATTCAGACAGACGGATCGGCTCCA 600
 QY 201 GluAlaG1uLeuG1nAsnCysLeuAlaG1nAlaG1YLeuG1nSerAspThrG1uCyThr 220
 Db 601 GAAGCAGATTGCAAAATGCTTGGCCGCGCTGCGCAAGTATCAGATGATCAGAAATGACC 660
 QY 221 SerAlaValAspLeuTrpSerLeuG1Y11eLeuLeuG1uMetPheSerG1YMetYs 240
 Db 661 TCAGCTGTGATCTGGAGGCTAGGAATCATTTTCTGGAATGTTCTCAGAAATGAA 720
 QY 241 LeuYnH1eThrValArgSerG1nG1uTrpYAlaAsnSerSerAla11e11eAspH1s 260
 Db 721 CTGAACATACAGTACAGATCTCAGGAATGGAAGCAAGTTCTGCTATTTATGATCAC 780
 QY 261 11ePheAlaSerYAlaValValAlaAla11eProAlaTYrH1sLeuArgAspLeu 280
 Db 781 AATATTGCGCAGTAAAGCATGTGTGAATGCCGAATTCAGCTTATCCTTAAGAGCCTT 840
 QY 281 11eYsSerMetLeuH1sAspAspProSerArgAlaG1Y11eProAlaG1uMetAlaLeuCyS 300
 Db 841 ATCAAAAGATCTTCATATATATCCAAAGCAAGATTCCTGCTGAATGCAATTTGCTC 900
 QY 301 SerProPheSer11eProPheAlaProH1s11eG1uAspLeuValMetLeuProThr 320
 Db 901 AGCCCATTTTATAGATTCCTTTGCGCCTCATATTGAAGATCTGGTATGCTCCACT 960
 QY 321 ProValLeuArgLeuLeuAsnVal11eAspAspAspTYrLeuG1YAsnG1uG1uTYr 340
 Db 961 CCAGTCTTAAGCTGTGATGTGCTGGATGATGATTATCTTGAGAAATGMAAGAGAAAT 1020
 QY 341 GluAspValValG1uAspVal1YsG1uG1uCyG1nYsTYrG1YProValValSerLeu 360
 Db 1021 GAAGATGTGTAAAGATGTAAAGAGAGGTCTCAAAATATGACCACTGTATCTTCTA 1080
 QY 361 LeuValProYAsG1uAsnPrG1YArgG1YG1nValPheValG1uTYrAlaAsnAlaG1Y 380
 Db 1081 CTTGTTCCAAAGAAATCTTGCAAGAGCAAGCTTTGTTGAATGCAAAATGCTGTT 1140
 QY 381 AspSerYValAlaG1nYsLeuLeuThrG1YArgMetPheAspG1YLYsPheValVal 400
 Db 1141 GATTCMAAGCTGCGAGAAATTAAGTACTGGAAGAGTGTGATGAGGAAAGTTGTGTG 1200
 QY 401 AlaThrPheTYrProLeuSerAlaTYrLYsAsG1YTYrLeuTYrG1nThrLeuLeu 419
 Db 1201 GCTACATTTACCGCTAGTGTCTTACAGAGGGGATATCTTATCAAACTTGGCTT 1257
 RESULT 4
 ABQ9383
 ID ABQ9383 standard; cDNA; 2008 BP.
 XX

AC ABQ99383;
 XX 25-FEB-2003 (first entry)
 XX Human coding sequence SEQ ID 116.
 DE Human; expressed sequence tag; EST; haematopoietic disorder;
 XX central nervous system disease; viral infection;
 XX peripheral nervous system disease; non-healing wound; infectious disease;
 XX immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 XX fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
 XX antileukemic; antineoplastic; immunosuppressive; neuroprotective;
 XX cytotoxic; haemostatic; virucide; antibacterial; fungicide;
 XX immunostimulant; cerebroprotective; gene therapy; gene; ss.
 OS Homo sapiens.
 XX WO200259260-A2.
 XX 01-AUG-2002.
 XX 16-NOV-2001; 2001MO-US042950.
 XX 17-NOV-2000; 2000US-00714936.
 XX (HYSB-) HYSBQ INC.
 XX Tang YF, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 XX Ren F, Xue AJ, Yang Y, Wehrman Y, Drmanac RT;
 XX WPI; 2002-590824/63.
 XX N-PSDB; ABP6497.
 XX New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.
 PT
 XX
 PS Claim 1; SEQ ID NO 116; 394pp; English.
 XX
 XX The present invention relates to novel human coding sequences (ABQ99268-
 CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotent or pluripotent state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/publicated_pct_sequences
 CC
 XX
 SQ Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.86e-208 Length: 2008
 Score: 2198.00 Matches: 418
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1

Query Match: 99.64% Indels: 0
 DB: 6 Gaps: 0
 US-10-798-532-2 (1-419) x ABQ99383 (1-2008)
 QY 1 MetAlaGlySerGlyCysAlaTrrGlyAlaGluProArgPheLeuGluAlaPheGly 20
 Db ATGGCGGGGATCCGGCTGCGCTGCGGGCGGAGCGCGCGCTTCTGGAGCGCTCGGG 138
 QY 21 ArgLeuTrrGlnValGlnSerArgLeuGlySerSerAlaSerAlaTyrArgVal 40
 Db CGGCTGTGCAAGTACAGAGCGGTGCGGTACCGCTCCGCTCGGTATGCGGTT 198
 QY 41 ArgCysCysGlyAspProGlySerProGlyAlaLeuGlyGlnPheLeuProGly 60
 Db CGCTGCTGCGGACCCCTGCTGCGCGCGCGCGCTTCAAGCAGTCTTCCGCGAGA 258
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTrrGlyPheArgGluAlaAlaLeu 80
 Db ACCACCGGGGCGCGGCTGCGCGCTGCGCGAGTATGTTCCGCAAGAAGAGCGCGCTG 318
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db GAACAGTTCAGAGTTCACAGAAACATCGACTTGTATGAGTGTTCATCAATCCACTT 378
 QY 101 SerProAsnValProSerArgCysLeuLeuGluLeuLeuAspValSerValSerGlu 120
 Db TCTCCAAATGTGCCATCAGCTGTCTGTGCTTGAATCTCGATGTCAGTGTTCGGA 438
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrrMetIleGlnHisCysAlaArg 140
 Db TTGCTCTTATATTCAGTCCAGCGGGTGTTCATGTGATGATACAGCATTCGCGCGA 498
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisIleGluGlyTyrValHisAlaAspLeuVal 160
 Db 499 GATGTTTGAAGCCCTTCTTCTTCATCAGAGGCTATATGTCATCGAGCTCAAA 558
 QY 161 ProArgAsnIleLeuTrrSerAlaGluAngGluCysPheLeuLeuIleAspPheGlyLeu 180
 Db 559 CACGTAACATATTTGGAGTGCAGAGTGAATGATTTTAACTCATTCATTGGAATT 618
 QY 181 SerPheLeuGluGlyAsnGlnAspValIleTyrIleGlnThrAspGlyTrrArgAlaPro 200
 Db 619 AGCTTAAAGAAAGGCAATCAGAGTGAATGATATTCAGACAGACGGGTATCGGCTCCA 678
 QY 201 GluAlaGluLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspPheGlnCysThr 220
 Db 679 GAAGCAGAAATTCGAAATTCCTTGGCCGAGCTGCGCTCAGAGTATCAGAAATGATCC 738
 QY 221 SerAlaValAspLeuTrrSerLeuGlyIleIleLeuGluLeuPheSerGlyMetLeuVal 240
 Db 739 TCAAGCTGTGATCTGGAGCTTGAAGATCTTTCAGAAATGTTCTCAGGAATGAA 798
 QY 241 LeuValHisThrValArgSerGlnGlyTrrPlyAlaAsnSerSerAlaIleLeuAspHis 260
 Db 799 CTGAACATACAGTCAATCAGTCAAGATGGAAGGCAACAGTCTCTCATTTATGATCAC 858
 QY 261 IlePheAlaSerIleValAlaValAlaAlaIleProAlaTyrHisIleLeuArgAlaPhe 280
 Db 859 ATATTTTCCAGTAAACAGATGGGATGCGCAATTCAGCTTATCACTTAAGAGACTT 918
 QY 281 IleValSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 Db 919 ATCAAAAGCATGCTTATGATGATCCAGACAGAAAGATTCCTGCTAAATGGCATGTGTC 978
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 979 AGCCCATTTCTTAGCATTCCTTTTGGCCCTCATATGAAGATCTGGTCAATGCTCCACT 1038
 QY 321 ProValLeuArgLeuLeuValLeuAlaAspAspTyrLeuGlyAsnGluGlnGlyTyr 340
 Db 1039 CCAAGTCTAAGATCTGTAATGTGATGATGATATCTTGAAGATGAAGAGAAATAT 1098

QY 341 GIUAAPValValIGIUAAPValIYSGIUGIUCYAGSGLNYSYRGIYProValValISerLeu 360
DB 1099 GAAGATGTTGTAGACATGTMAAAGAGAGTGTCAAAAATATGACACAGTGTATCTCTA 1158
QY 361 LeuValProLYSGIUGIUNProGIYARGIYGLINValPheValIGIUNYRValAaenValAGIY 380
DB 1159 CTTGTTCCAAAGGAAATCTGGCAGAGCAAGTCTTGTGTAAGTAAATGCTGGT 1218
QY 381 AepSerValAaIaIaGLNYSLeuThRGlyArgMetPheAepGIYLYSPheValVal 400
DB 1219 GATTCGAAAGCTCCGCAAAATTAAGTACGTGAGAGATGTTTATGGAAATTTGTTGTG 1278
QY 401 AATThPheTYrProLeuSerAlaTYrLYSARGIYTYrLeuTYrGLNThLeuLeu 419
DB 1279 GCTACATCTACCCGCTGAGTGCCTACAGAGGGGATATCTGTCAAACTTGCTT 1335
RESULT 5
AAI66828
ID AAI66828 standard; cDNA; 2622 BP.
AC AAI66828;
XX
XX 07-JAN-2002 (first entry)
DT Human protein kinase polypeptide 13887 encoding cDNA.
DE
XX
XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KM cytostatic; antiinflammatory; immunosuppressive; cardiac; hepatocrophic;
KW gene therapy; 88.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT 5'UTR 1..266
FT CDS /tag= a
FT /tag= b 267..1529
FT /product= "protein kinase 13887"
FT /note= "coding sequence specifically claimed"
FT 3'UTR 1530..2622
FT /tag= c
XX
XX WO200173050-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001MO-US009483.
XX
XX 24-MAR-2000; 2000US-0191846P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers R;
XX
XX WPI, 2001-611632/70.
XX
XX P-PSDB; AAG65767.
XX
XX
XX New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
XX
XX PT useful in diagnosis of cancer or cellular proliferation or
XX
XX PT differentiation disorders and to screen for polypeptide modulators useful
XX
XX PT to treat such conditions.
XX
XX
XX Claim 1; Fig 10A-C; 169pp; English.
XX
XX
XX The invention provides novel human protein kinase polypeptides, 3714,
XX
XX 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
XX
XX protein kinase polypeptides can be expressed by standard recombinant
XX
XX methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
XX
XX are useful for diagnostic and screening methods to identify subjects (at
XX
XX risk of) having cancer or cellular proliferation and/or differentiation
XX
XX disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
XX
XX modulators are useful for the treatment of cancer, particularly colon
XX
XX cancer or cellular proliferation and/or differentiation disorders. Other

CC disorders associated with 3714, 16742, 23546 or 13887 expression or
CC activity that can be treated include bone related disorders, inflammatory
CC disorders, autoimmune diseases, cardiovascular disorders and liver
CC diseases. The present sequence represents a human protein kinase
CC polypeptide 13887 encoding cDNA
XX
SQ Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 8, 67e-208 Length: 2622
Score: 2198.00 Matches: 418
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.64% Indels: 0
DB: 4 Gaps: 0
US-10-798-532-2 (1-419) x AAI66828 (1-2622)
QY 1 MetAlaGLYSerGIYCYaAlaTRPGLYAlaGLUPROProArgPheLeuGLUAlaPheGIY 20
DB 267 ATGGCGGATCCGCGCTGCGCCCTGGGGCGCGAGCCGCCGTTTCTGGAGGCTTCGGG 326
QY 21 ArgLeuTRPGLNValGlnSerArgLeuGLYSerGIYSerSerAlaSerValTYrArgVal 40
DB 327 CGGCTGGAGAGTACAGACCGCTCTGGGTAGCGGCTCTCCGCTCGATGTCCGGTT 386
QY 41 ArgCYaGLYASnProGIYSerProProGIYAlaLeuLYSGINPheLeuProProGIY 60
DB 387 CGGTGTGGGGAACCTTGCTGCGCCCGCGCGCCTCAACAGATTTGGCGCAGGA 446
QY 61 ThrThrGLYAlaAlaAlaSerAlaAlaGLIYRGIYRGIYRGIYRGIYRGIYRGIYRGIYR 80
DB 447 ACCACCGGGGCTGGCGCCCTGCGCCCGCGAGTGTGTTCCGAAAGAGGGCGGCGCTG 506
QY 81 GluGlnLeuGlnGLYHISArgAsnILEValThrLeuTYRGIYValPheThrILEHisPhe 100
DB 507 GAACAGTGCAGGGTACAGAAACATGTAAGTGTGTATGAGTGTTCATTCACATCCACTT 566
QY 101 SerProAsnValProSerArgCYaLeuLeuGLULeuLeuAspValSerValSerGIU 120
DB 567 TTCGCAATGTGCATCAAGCTGTCTGTGAACCTCTGGATGTCACTGTTCGGA 626
QY 121 LeuLeuLeuTYrSerSerHISGlnGLYCYSerMetTrpMetILEGNIHISCYaAlaArg 140
DB 627 TTGCTCTTATATCCAGTACACAGGGTGTTCATGTGATGATACAGCATTTGTCCGA 686
QY 141 AapValLeuGLUAlaLeuAlaPheLeuHISGLIYTYrValHISAlaAspLeuLYS 160
DB 687 GATGTTTGAAGCCCTTCTTTCTTCATCAATGAGGCTATGTCCATCCGACCTCAAA 746
QY 161 ProArgAsnILEuTrpSerAlaGluAengIUCYSPheLYLeuILEaPheGLIYLeu 180
DB 747 CCACGTAACATATTGGAGTGCAGAGATGATGATTTAACTATTCATTCATTGGACTT 806
QY 181 SerPheLYSGIUYASnGlnAspValLYSTYrILEGINTHAspGLIYTYrArgAlaPro 200
DB 807 AGCTTCMAAGAAAGGCATCAGATGTAAAGTATTTCACACAGACGGGATCGGCTCCA 866
QY 201 GIUAIGIUGIUGIAsnCYaLeuAlaGlnAlaGLIYLeuGlnSerAspThRGIUCYThr 220
DB 867 GAAGCAGAAATGCAAAATTTGCTTGCCCAAGGCTGGCTCAAGATGATACAGAAATTCAC 926
QY 221 SerAlaValaAspLeuTrpSerLeuGIYILELeuLeuGLUmePheSerGIYMeLYS 240
DB 927 TCAGCTGTGATGTGTGAGAGCTTGAAGAACATTTTAACTGGAATGTTCACAGAAATGAA 986
QY 241 LeuLYSHISThrValArgSerGlnGLUTRPLYSAlaAsnSerSerAlaILEILEaPHis 260
DB 987 CTGAACATATACAGTCAAGATCTCAGAGATGAAAGGCAAGTTCGTCTTATTATGATCAC 1046
QY 261 IlePheAlaSerLYaAlaValAlaAsnAlaAlaIleProAlaTYrPHISLeuArgAspLeu 280
DB 1047 ATATTTCAGTAAGCAAGTGTGTGAATGCCGCAATTCAGCTTATCATTAAAGACCTT 1106

QY 281 IleySerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 1107 ATCAAAAGCATGCTCATATGATGATCCAGAGAGATTCCTGCTGATAATGGCATTGTGC 1166
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 1167 ACCCATTTCTTATAGCATTCCTTTTGGCCCTCATATTGAAGATCTGGTCAATGGTCCCACT 1226
 QY 321 ProValLeuArgLeuLeuValLeuAspAspAspPtyrLeuGluYasnGluGluGluTyr 340
 DB 1227 CCAGTGTCTAAGACTGCTGATGTGCTGATGATGATTAATCTTGAGATGAATGAAGAGATAT 1286
 QY 341 GluAspValValGluAspValIlyssGluGluCysGluIlyeTyrGlyProValValSerLeu 360
 DB 1287 GAAGATGTTGTGAGATGTAAGAAGAGATGTCAAAATAATGACCACTGATTTCTCA 1346
 QY 361 LeuValProIlysgluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 DB 1347 CTTGTTCAAAGGAATATCTGCGCAGAGCAAGCTTTTGTGATGATCAATGCTGGT 1406
 QY 381 AspSerIlyAlaAlaGluIlyleuLeuThrGlyArgMetPheAspGlyIlyspPheValVal 400
 DB 1407 GATTCCAAAGCTGCGCAGAAATTAAGTGAAGATGTTGATGGAAGTTGTGTGTG 1466
 QY 401 AlaThrPheTyrProLeuSerAlaTyrIlyArgGlyTyrIleuTyrGlnThrLeuLeu 419
 DB 1467 GCTACATTCACCCGCTGAGTGTCAAGAGGGGATCTGTATCAAACTTGTCTT 1523
 RESULT 6
 AAH72714
 ID AAH72714 standard; cDNA; 4065 BP.
 AC AAH72714;
 XX 19-SEP-2001 (first entry)
 DT Human cervical cancer marker nucleic acid 3988.
 XX
 KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200142467-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000WO-US033312.
 XX
 PR 08-DEC-1999; 99US-0169681P.
 PR 21-DEC-1999; 99US-0171350P.
 PR 14-MAR-2000; 2000US-0189315P.
 PR 12-MAY-2000; 2000US-0203791P.
 PR 09-JUN-2000; 2000US-0210600P.
 PR 21-JUL-2000; 2000US-0220114P.
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Deeds J, Berger A, Zhao X;
 XX
 DR WPI; 2001-375006/39.
 XX
 PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 CC for assessing and detecting compounds for treating the cancer.
 XX
 PS Claim 1; Page 810-811; 1051pp; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for

CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy
 XX
 SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;
 Alignment Scores:
 Pred. No.: 1,656-207 Length: 4065
 Score: 2198.00 Matches: 418
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.64% Indels: 0
 DB: 4 Gaps: 0
 US-10-798-532-2 (1-419) x AAH72714 (1-4065)
 QY 1 MetAlaGlySerGlyCysAlaThrProGlyValAlaGluProProArgPheLeuGluAlaPheGly 20
 DB 268 ATGGCGGGATCCGGCTGGCGCTGGGGCGGAGCGCGCTTTCTGAGGCTTCGGG 327
 QY 21 ArgLeuThrGlnValAlaGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB 328 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCTCCGCTCGGTATGGGT 387
 QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuIlysglnPheLeuProProGly 60
 DB 388 CGCTGTGGCGCAACCTGGCTGGCCCGCGCGCTCAAGACAGTTCCTTGGCCGACAGA 447
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
 DB 448 ACCACCGGGGCTGGCGCTCTGCGCCCGCGAGTATGGTTCCGCAAGAGAGGGCGCGCTG 507
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB 508 GAACGTTGACAGGTGACAGAAACATCTGATCTTGTATGATGATGTTTAACTTCACTTT 567
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuAspValSerValSerGlu 120
 DB 568 TCTCCAAATGTCCATCAAGCTGTCTGTGAACCTCGATGATGATGATGATGATGATGAT 627
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTyrMetIleGlnHisCysAlaArg 140
 DB 628 TTGCTCTTATATTCAGTCCACAGAGGTTGTTCATGTGATGATGATGATGATGATGATGAT 687
 QY 141 AspValIleuGluAlaLeuAlaPheLeuHisIleGluGlyTyrValHisAlaAspLeuIly 160
 DB 688 GATGTTTGGAGCCCTTGCTTTCTTCAATGAGGCTATGTCATCGAATCGAATCTCAAA 747
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheIlyLeuIleAspPheGlyLeu 180
 DB 748 CCAAGTAAATATTTGGAGTGCAGAGATGAATGTTTAACTCATTCATTCATTCATTCATTC 807
 QY 181 SerPheIlyGluGlyAsnGlnAspValIlyTyrIleGlnThrAspGlyTyrArgAlaPro 200
 DB 808 ACCTTCAAGAGAGCAATCAGATGAAGTATATTCAGACAGAGCGGATATCGGGCTCCA 867
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyIlyeGlnSerAspThrGluCysThr 220
 DB 868 GAAGCAGAAATTCAAATATGCTTGGCCAGGCTGGCTGCAGAGATGATACAAATGTACC 927
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetIly 240
 DB 928 TCAAGCTGTGATCTGTGGAGCTTAGAATCATTTTAACTGGAATGTCTCAAGAAATGAAA 987
 QY 241 LeuIlyHisThrValArgSerGlnGluTyrIlyAlaAsnSerSerAlaIleIleAspHis 260
 DB 988 CTGAAGCATACAGTCAATCTCAGAGATGAGAGCAACAGTTCGTATTTATGATCAC 1047
 QY 261 IlePheAlaSerIlyAlaValAlaAsnAlaIleProAlaTyrHisIleuArgAspLeu 280
 DB 1048 ATATTTGCCAGTAAAGCAATGTGTAATGCCGAATTCAGCTTATCACTTAAGAGACTT 1107
 QY 281 IleySerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

Db 1108 ATCAAAAGCATGCTTCATGATGATCCAGAGAAATTCCTGCTGAATGCAATTGTGC 1167
 QY 301 SerProPheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 1168 AGCCCATTTCTTATAGCATTCCTTTTGGCCCTCATATTGAAGATCTGGTCATGCTTCCACT 1227
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspSerTyrLeuGluAsnGluGluTyr 340
 Db 1228 CCAGTCTAAGCATGCTGGAATGCTGGAGATGATGATATCTTCAAGATGAAGAGAAATAT 1287
 QY 341 GluAspValValGluAspValValGluGluGluGluGluGluGluGluGluGluGlu 360
 Db 1288 GAAGAGATGTTGTAGAGATGTTAAAGAGAGATGTTCAAAAATATGACCCAGTGTATCTTA 1347
 QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 Db 1348 CTGTGTTCCAAAGAGAAATCTGTGCAAGGACAAAGTCTTTTGTAGTATGAAATGCTGGT 1407
 QY 381 AspSerIleAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 Db 1408 GATTCCAAAGCTCGCGAATTAATGACTGGAAGAGATGTTGATGGAAATGTTGTGTG 1467
 QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 Db 1468 GCTACATTTCTACCGCTGAGTGCCTACAGAGAGGATATCTGTATCAAACTTGTCTT 1524

RESULT 7

ID ADL45238
 ADL45238 standard; DNA; 4065 BP.

ADL45238;

20-MAY-2004 (first entry)

Human ovarian cancer DNA marker #19128.

Human; ovarian cancer; de; tumour; cytostatic; DNA marker.

Homo sapiens.

MO200170979-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US009126.

21-MAR-2000; 2000US-0191031P.

25-MAY-2000; 2000US-0207124P.

15-JUN-2000; 2000US-0211940P.

07-JUL-2000; 2000US-0216820P.

25-JUL-2000; 2000US-0220661P.

21-DEC-2000; 2000US-0257672P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Lee J, Little J;

WPI; 2001-611502/70.

Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.

Disclosure; SEQ ID NO 19128; 106bp; English.

The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of

CC the patient an antisense oligonucleotide complementary to a marker of the
 CC invention. The markers are useful for assessing if a patient is afflicted
 CC with ovarian cancer, which involves comparing the level of expression of
 CC a marker in a patient sample and a normal level of expression of the
 CC marker in a control non-ovarian cancer sample. A difference between the
 CC expression levels indicates ovarian cancer. The level of expression of a
 CC marker corresponds to a secreted protein or to a transcribed
 CC polynucleotide or its portion. The level of expression of the marker is
 CC assessed by detecting the presence in the sample, a protein or protein
 CC fragment corresponding to the marker. The presence of protein or protein
 CC fragment is detected using an antibody that specifically binds with the
 CC protein or protein fragment. Alternatively, the level of expression of
 CC the marker is assessed by detecting the presence of a transcribed
 CC polynucleotide which anneals with the marker or anneals with a portion of
 CC the polynucleotide comprising the marker, under stringent conditions. The
 CC marker is also used for monitoring the progression of ovarian cancer in a
 CC patient which involves detecting expression of the marker in a patient
 CC sample at a first point in time, repeating the method at a subsequent
 CC time and comparing the level of expression. The method is carried out
 CC using an ovarian tissue sample. A composition comprising a marker,
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.
 CC This sequence represents a human ovarian cancer DNA marker of the
 CC invention.

Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.: 1,65e-207 Length: 4065
 Score: 2198.00 Matches: 418
 Percent Similarity: 99.764 Conservative: 0
 Best Local Similarity: 99.764 Mismatches: 1
 Query Match: 99.648 Indels: 0
 DB: 5 Gaps: 0

US-10-798-532-2 (1-419) x ADL45238 (1-4065)

QY 1 MetAlaGlySerGlyCysAlaIleTrrGlyAlaGluProProArgPheLeuGluAlaPheGly 20
 Db 268 ATGGCGGAGATCCGGCTGCGCTGGGGCGGAGCCGCCGTTTCTTGAGGCTTCGGG 327
 QY 21 ArgLeuTrrGlyValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 328 CGGCTGTGGCAGGATCAGAGCCGCTCTGGGTACGGCTCTCCGCTCGGTATGCGGT 387
 QY 41 ArgCysGlyAsnProGlySerProGlyAlaLeuLysGlnPheLeuProGly 60
 Db 388 CGCTGTGGCGCAACCTGTGCGCCCGCGCCCTCAACAGATCTTGGCCGACAGA 447
 QY 61 ThrThrGlyAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaLeu 80
 Db 448 ACCACCGGGGCTGGCGCTCTGCCCGCGATGTTCCGAAAGAGGGCGCGCTG 507
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 508 GAACAGTTTCAAGGATCAGAAACATCGATCTTGTATGAGAGTGTTCATTCACCTTT 567
 QY 101 SerProAsnValProSerArgCysLeuLeuGluLeuLeuAspValSerValSerGlu 120
 Db 568 TTCCAAATGTGCATCAGCTGTCTGTGTTGAACCTCGATGATCAGTGTTCGGA 627
 QY 121 LeuLeuLeuTyrSerSerIleGlnGlyCysSerMetTrrMetIleGlnHisCysAlaArg 140
 Db 628 TTGCTCTTATATTCAGTACACAGGGTGTTCATGTGATATATACGATTTGCCCCA 687
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisIleGluGlyTyrValHisAlaAspLeuLys 160
 Db 688 GATGTTTGAAGGCCCTTCTTCTTCATCATGAGGGTATATGTCAGGACCTCAAA 747
 QY 161 ProArgAsnIleLeuTrrSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
 Db 748 CCACGTAATATTTGGAGTGACAGAGATGATGATTTAACTCATTAATCTTGGACTT 807
 QY 181 SerPheLysGluGluGlnAsnIleAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200

```

DB 808 AGCTTCAAGAGGACATCAGATGTAATATATTCAGACGAGCGGATATGGGCTCCA 867
QY 201 GUAAGAGLUGLUGLanGysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
DB 868 GAAGCGAATATGGCAAAATTCCTGGCCGAGCTGGCTGAGATGATCAGAAATAC 927
QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuGlnMetPheSerGlyMetCys 240
DB 928 TCAAGCTTGTATGTGAGGCTTGAAGAAATATTTACTGAAATGTTCTCAGAAATGAAA 987
QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
DB 988 CTGAACATACATCAGTCAATCTCGAGATGAGAGCAAAACAGTTCCTGATTAATGATCAC 1047
QY 261 IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTrpHisLeuArgAspLeu 280
DB 1048 AATATTGCCAGTAAACAGTGGTGAATGCGCAATTCAGCCTATCAGCTAAGAGACCTT 1107
QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuCys 300
DB 1108 ATCAAAAGCATGCTTATGATGATGATCCAGCAGAAAGAAATTCCTGTAATGGCATTGTGC 1167
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB 1168 AGCCCATTCCTTAAGCATTCCTTTGGCCCTCATATTTGAAGATGTGCTCATGCTCCACT 1227
QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTrpLeuGlyValGlnGluGluTrp 340
DB 1228 CCAGTCTAAGACTGCTGAAATGCTGATGATGATTAATCTTGAGAAATGAGAGAAATAT 1287
QY 341 GluAspValValGluAspValLysGluGluGlnGlySerGlyTrpValProValIleSerLeu 360
DB 1288 GAAGATGTTGTAGAAATGTAAGAGAGGTGCAAAATATGACACAGGATCTCTA 1347
QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTrpAlaAsnAlaGly 380
DB 1348 CTGTGTTCCAAAGAAAATCTGGCAGAGCAAGATCTTTCTTGAGATGCAAAATGCTGGT 1407
QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyValGlyMetPheAspGlyLysPheValVal 400
DB 1408 GATTCCAAAGCTGCGCAAAATTAAGTGAAGATGTTGATGGAAGTTGTTGGTG 1467
QY 401 AlaThrPheTrpProLeuSerAlaIleLysArgGlyIleLysTrpLeuLeuLeuLeu 419
DB 1468 GCTACATTCACCCGCTGAGTGCCTACAGAGGGGATATCTGATCAAACTTGCTT 1524

```

RESULT 8
AAS06734
ID AAS06734 standard; cDNA; 1260 BP.
XX AAS06734;
AC 12-SEP-2001 (first entry)
DT 12-SEP-2001 (first entry)
XX
DB Polynucleotide sequence encoding human protein kinase #34.
XX
XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
OS Homo sapiens.
XX
XX WO200138503-A2.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000MO-US032085.
XX
XX 24-NOV-1999; 99US-0167482P.
XX
XX (SUB-) SUGEN INC.

```

XX Plozman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX WPI; 2001-343950/36.
DR P-PSDB; AAU03534.
XX
PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections.
XX
PS Example 1; Fig 1; 433bp; English.
XX
CC AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
XX
SQ Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;

```

Alignment Scores:
Pred. No.: 1,476-207 Length: 1260
Score: 2191.00 Matches: 417
Percent Similarity: 99.52% Conservative: 0
Best Local Similarity: 99.52% Mismatches: 2
Query Match: 99.32% Indels: 0
DB: Gaps: 0

US-10-798-532-2 (1-419) x AAS06734 (1-1260)

```

QY 1 MetAlaGlySerGlyCysAlaIleTrpGlyValGluProProArgPheLeuGlnAlaPheGly 20
DB 1 ATGGGGGAGTCCGGCTGCGCTGGGCGCGGAGCGCGGCTTCTGAGGCTTCGGG 60
QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyraArgVal 40
DB 61 CGGCTGTGGCAGGTACAGAGCGCTGTGGTACGGCTCTCCGCTCGGTATCGGTT 120
QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuLysGlnPheLeuProProGly 60
DB 121 CGCTGCTGGGCAACCTTGCTCGCCCCCGGCGCCCTCAAGCAATGTTCTTGGCCGAGA 180
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTrpGlyPheAspArgLysGluAlaLeu 80
DB 181 ACCACCGGGGCTGGCGCTCTGCGCCGAGATGATGTTCTCCGAAAGAGGGCGCGCTG 240
QY 81 GluGlnLeuGlnGlnHisArgAsnIleValThrLeuTrpGlyValPheThrIleHisPhe 100
DB 241 GAACAGTTCAGAGGTCAAGAAACATCGTGAATCTTTGATGAGATGTTTCAATCCACTT 300
QY 101 SerProAsnValProSerArgCysLeuLeuGlnLeuLeuAspValSerValSerGlu 120
DB 301 TCTCCAAATGTCCATCAAGCTGTGCTGTTGAATCTCGAGATGATGATGTTTCCGAA 360
QY 121 LeuLeuLeuTrpSerSerHisGlnGlnCysSerMetTrpMetIleGlnHisCysAlaArg 140
DB 361 TTGCTTTATATTCAGTCAACAGGATGTTTCAATGATGATGATGATGATGATGATGATGAT 420
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGluGlyTrpValHisAlaAspLeuLys 160

```

Db 421 GATGTTTGGAGGCCCTTGCTTTCTTCATCATAGAGGCTAGTCCATGCGGACCTCAAA 480
 Qy 161 ProArgAsnIleLeuTrpSerAlaGluAenGluCysPheLeuIleAspPheGlyLeu 180
 Db 481 CCACGTRACATATTTGGAGTGCAGAGATGAATGTTTAACTCATTTGACTTGGACTT 540
 Qy 181 SerPheLeuGluGlyAenGluAenValIysTyrIleGlnThrAspGlyTyrArgAlaPro 200
 Db 541 AGCTTCAAAGAGGCAATCAGATGTAAATATATTCAGACAAACGGGTATCGGGCTCCA 600
 Qy 201 GluAlaGluLeuGlnAenCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 Db 601 GAAGCGAAATGCAAAATGCTTGCGCCAGGCTGCGCTGAGGTATATACAGATGTACC 660
 Qy 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
 Db 661 TCAGCTGTTGATGTGGAGCCAGAGATCATTTTCTGGAATGTTCTCAGGAATGAAA 720
 Qy 241 LeuLysHisThrValIysSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
 Db 721 CTGAAACATACAGTCAATCTCAGGAATGGAAGCAAAACAGTTCCTCATTTATTGATCAC 780
 Qy 261 IlePheAlaSerIysAlaValIleAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
 Db 781 ATATTGGCCAGTAAAGAGTGTGAATGCCGCAATTCAGCTTATACCTAAGAGACCTT 840
 Qy 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 Db 841 ATCAAAAGCATGCTTCATGATGATCCAGAGAAAGAAATTCCTGTAATGGCATTTGTGC 900
 Qy 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 901 AGCCCATCTCTTATGACATTCCTTTGCCCCCTCATATGAAGATCTGTCATGCTCCACT 960
 Qy 321 ProValLeuArgLeuLeuAenValIleAspAspAspTyrLeuGluAenGluGluTyr 340
 Db 961 CAGTCTTAAGACTGCTGAATGTGCTGAGATGATATATCTTGAAGATGAAGGAAATAT 1020
 Qy 341 GluAspValIleGluAspValIysGluGluCysGlnLysTyrIleProValIleSerLeu 360
 Db 1021 GAAGATGTTGTAAAGATGTAAAGAGGAGTGTCAAAAATATGACACAGTGGTATCTCA 1080
 Qy 361 LeuValProLysGluAenProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 Db 1081 CTGTGTCCAAAGGGAATCTGCGCAAGGACAAAGTCTTGTTCAGATGCAATGCAATGCTGT 1140
 Qy 381 AspSerValAlaIleGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValIle 400
 Db 1141 GATTCGAAGCTGCGCAAAATTAAGTGAAGGAGTGTGTGATGGAAGTGTGTGTG 1200
 Qy 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 Db 1201 GCACATTTCTACCGGTGAGTGTCTCAAGAGGGGATATCTGTATCAAACTTGCTT 1257
 RESULT 9
 AAD18818
 ID AAD18818 standard; cDNA; 1824 BP.
 AC AAD18818;
 XX 18-DEC-2001 (first entry)
 DT 18-DEC-2001 (first entry)
 XX Human kinase (PKIN) -3 cDNA.
 DE Human kinase: PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
 XX cancer; allergy; sarcoma; leukemia; acquired immune deficiency syndrome;
 KM AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
 KM atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
 KM myasthenia gravis; cirrhosis; cataract; growth and development disorder;
 KM seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
 KM lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
 KM asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
 KM antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.

XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 169..1428
 FT /tag=8
 FT /product= "Human PKIN-3 protein"
 XX
 XX PD WO200181555-A2.
 XX
 XX 01-NOV-2001.
 XX
 XX PF 20-APR-2001; 2001WO-US012992.
 XX
 XX 20-APR-2000; 2000US-0199021P.
 PR 28-APR-2000; 2000US-0200226P.
 PR 05-MAY-2000; 2000US-0202339P.
 PR 11-MAY-2000; 2000US-0203505P.
 PR 18-MAY-2000; 2000US-0205564P.
 PR 26-MAY-2000; 2000US-0207739P.
 PR 01-JUN-2000; 2000US-0208795P.
 XX
 XX PA (INCYTE GENOMICS INC.
 XX
 XX PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
 PI Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
 PI Paterson C, Burrill JD, Marcus GA, Zingler KA, Reclon SA, Lu Y;
 PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
 PI Walsh RP, Rameshwar J, Borowsky ML, Au-Young J, Hillman JL;
 PI Gururajan R;
 XX
 XX DR WPI; 2001-611740/70.
 DR P-PSDB; AAE11769.
 XX
 XX PT Human kinases and nucleic acids, useful for preventing diagnosing and
 PT treating cancers, inflammation and immune disorders.
 XX
 XX PS Claim 5; Page 152; 166pp; English.
 XX
 XX The present invention relates to human kinases (PKIN) and the nucleic
 CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
 CC used in the prevention, diagnosis and treatment of diseases cancers,
 CC adenocarcinoma, leukemia, sarcoma, immune disorder, Addison's disease,
 CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
 CC gout, microbial infections, cardiovascular disease and/or inflammation,
 CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
 CC infarction, cataract, growth and development disorder, seizure disorder,
 CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
 CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
 CC PKIN may be used to treat disorders associated with decreased PKIN
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of PKIN by expressing inactive proteins or to
 CC supplement the patient's own production of PKIN. PKIN nucleic acids may be
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
 CC a host cell and culturing the cell to express the protein. PKIN nucleic
 CC acid and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acid sequences in samples and therefore which patients may be in
 CC need of restorative therapy. The present sequence is human PKIN-3 cDNA
 XX
 XX SQ Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,536-207 Length: 1824
 Score: 2191.00 Matches: 417
 Percent Similarity: 99.52% Conservative: 0
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.32% Indels: 0
 DB: 4 Gaps: 0
 US-10-798-532-2 (1-419) x AAD18818 (1-1824)
 Qy 1 MetaGlySerGlyCysAlaTrpGlyAlaGluProPheArgPheLeuGluAlaPheGly 20

```

Db      169 ATGGCGGAGATCCGGCTGCGCTGGGCGCGAGCCGCGGCTTTCTCGAGGCGCTTCGCG
Qy      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      229 CGGCTGTGGCAGGTACAGAGCCGCTCTGGTAGCGGCTCTCCGCTCGGAGTATCGGGTT 288
Qy      41 ArgCysCysGlyAsnProGlySerProProGlyValAlaLeuGlnInheLeuProProGly 60
Db      289 CGCTGTGGCGCAACCTGTGCTGCGCCCGCGGCGCTCAAGACATCTTGTGCGCAGGA 348
Qy      61 ThrThGlyValAlaAlaSerAlaAlaGluTyrGlyPheArgPheGluArgAlaAlaLeu 80
Db      349 ACCACCGGGGCTCGGCGCTTCGCGCGAGTATGTTTCGCAAAAGAGGGCGGCGCTG 408
Qy      81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db      409 GAACAGTTGGCAGGGTACAGAAACATCGTGAATTTGATGAGTGTTCATCAATCCACTTT 468
Qy      101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
Db      469 TCTCCAAAGTGTGCATCACGCTGTGTGCTTGAACTCTGTGATGTCAGTGTTCGGA 528
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db      529 TTGCTCTTATATTCAGTCCAGGCTTCCATGTGATGATGATGATGATGATGATGATGATG 588
Qy      141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuLys 160
Db      589 GATGTTTGGAGGCGCTTGTGCTTCTTCATCATGAGGCTATGCTCATGCGAGCTCAAA 648
Qy      161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
Db      649 CCACGTAACATATGTGGAGTGCAGAGAAATGTTTAACTCATGTGACTTGACCTT 708
Qy      181 SerPheLysGlnGlyAsnGlnAspValLysTyrCileGlnThrAspGlyTyrArgAlaPro 200
Db      709 AGCTTCAAGAGAGGCAATCGAGATGTAATATTCAGACAGCGGATTCGGGCTCCA 768
Qy      201 GluAlaGlnLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspTrpGluCysThr 220
Db      769 GAAGCAGAAATTCGAAATTCCTTGCCAGGCTGCGCTGACAGATGATACAGAAATGACC 828
Qy      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys 240
Db      829 TCGAGCTGTGATGTGGAGCCCAAGGAAATCTTCTGAAATGTTCTCAGAAATGAAA 888
Qy      241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
Db      889 CTGAAACATACATCAGATCTTCAGAAATGAGGCAAAACGTTCTGCTATTATTGATCAC 948
Qy      261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgPheLeu 280
Db      949 ATATTGCGCAGTAAGCAGTGTGAATGCGCAATTCACGCTATACCTTAAGAGACCTT 1008
Qy      281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuCys 300
Db      1009 ATCAAAAGCAATGCTTATGATGATCCAAAGCAAGAAATTCCTCTGAAAGGCAATGTGC 1068
Qy      301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db      1069 AGCCCATCTTTTACATTCCTTTTGCCTCATATTGAAAGATCTGTGATCTTCCACT 1128
Qy      321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlnGlnGlnGlnTyr 340
Db      1129 CCAAGTCTTAAGACTGCGTGAATGCTGTGATGATGATATATCTTGAAGATGAAGGAAATAT 1188
Qy      341 GluAspValValGluAspValLysGlnGluCysGlnLysTyrGlyProValIleSerLeu 360
Db      1189 GAAGAGTGTGTAAAGATGTAAGAGAGAGATGCAAAATAATGACACAGTGTATCTCTA 1248
Qy      361 LeuValProLysGlnAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

```

```

Db      1249 CTTGTTCCAAAGGAATTCCTGCGACAGGACAAATCTTTGTTGATGATCAAAATGCTGCT 1308
Qy      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db      1309 GATTCCAAGAGCTGCCAGAAATTAATCACTGCAAGAGATGTTGATGGAAAGTTTGTG 1368
Qy      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db      1369 GCTACATTCACCGCTGAGTGCCTTACAGAGGGAGATCTGTATCAAACTTGCCTT 1425

RESULT 10
ADBS2857
ID ADBS2857 standard; DNA; 3244 BP.
AC ADBS2857;
XX
XX 04-DEC-2003 (first entry)
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3399.
XX
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; db.
XX
XX Rattus norvegicus.
XX
XX MO2003065993-A2.
XX
XX 14-AUG-2003.
XX
XX 04-FEB-2003; 2003WC-US003482.
PF
XX
XX 04-FEB-2002; 2002US-0353171P.
XX
XX 13-MAR-2002; 2002US-0363534P.
XX
XX 08-APR-2002; 2002US-0370248P.
XX
XX 10-APR-2002; 2002US-0371134P.
XX
XX 10-APR-2002; 2002US-0371135P.
XX
XX 10-APR-2002; 2002US-0371150P.
XX
XX 11-APR-2002; 2002US-0371413P.
XX
XX 19-APR-2002; 2002US-0373601P.
XX
XX 22-APR-2002; 2002US-0374139P.
XX
XX 08-MAY-2002; 2002US-0378370P.
XX
XX 09-MAY-2002; 2002US-0378652P.
XX
XX 09-MAY-2002; 2002US-0378653P.
XX
XX 09-MAY-2002; 2002US-0378655P.
XX
XX 09-JUL-2002; 2002US-0394230P.
XX
XX 09-JUL-2002; 2002US-0394253P.
XX
XX 04-SEP-2002; 2002US-0407688P.
XX
XX 28-JUN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
PA
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
PI
XX
XX WPI, 2003-731472/69.
XX
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX Claim 44; SEQ ID NO 3399; 874bp; English.
XX
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
XX preparing a gene expression profile of a tissue or cell sample exposed to
XX the compound, and comparing the gene expression profile to a database
XX comprising data or information on the Tox mean and non-Tox mean value.
XX The method is useful for predicting or identifying at least one toxic
XX effect, particularly hepatotoxicity, of a test or unknown compound. The
XX genes listed in the specification are useful as diagnostic or toxicity

```


CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.

XX Sequence 3244 BP; 800 A; 710 C; 819 G; 915 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 1 84e-206 | Length: | 3244 |
|------------------------|-----------|---------------|------|
| Score: | 2186.00 | Matches: | 415 |
| Percent Similarity: | 99.52% | Conservative: | 2 |
| Best Local Similarity: | 99.05% | Mismatches: | 2 |
| Query Match: | 99.09% | Indels: | 0 |
| DB: | 10 | Gaps: | 0 |

US-10-798-532-2 (1-419) x ADB52857 (1-3244)

```
QY 1 MetAlAGlySerGlyCysAlaTRPGlyAlaGluProPheLeuGluAlaPheGly 20
DB 162 ATGGCCGGGTCGGCTGCGCGTGGGGGCGGAGCCGCGCTTCTCGAGGCTTCGGG 221
QY 21 ArgLeuTRPGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArgVal 40
DB 222 CGGCTGTGGCAGGTACAGAGCCGCTTAGGCGAGCGGCTCTCGGCTCGGTGACCGGGTG 281
QY 41 ArgCysGAGlyAsnProGlySerProPheGlyAlaLeuGlyGlnPheLeuProPheGly 60
DB 282 CGGTGCTGGCGCACTCAGAGCTGCGCCCGCGGCTCTCAAGAGTTCCTCGCTCCGGGA 341
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaAlaLeu 80
DB 342 ACCACCGGGGCTCCGCTGCGCTGCGAGTATGTTTCGAAAGAGAGGCGAGCGCTG 401
QY 81 GluGlnLeuGlnGlnYHisArgAsnIleValThrLeuTYrGlyValaPheThrIleHisPhe 100
DB 402 GAGCAGATTGCGAGGTCACAGAGAACATCGTACTTTATACGAGCTCTTACCATACACTTC 461
QY 101 SerProAsnValProSerArgCysLeuLeuGluGlnLeuAspValSerValSerGlu 120
DB 462 TTCCTCAAAATGTGCATCACGCTCTCTGTTGCTTGAACCTCTGATGTCAGCGTTTGGAA 521
QY 121 LeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
DB 522 TTGCTGTATATTCAGTCCAGGAGTGTCTCATGATGATCCAGCATGTCGCA 581
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisIleGluGlyTYrValHisAlaAspLeuVal 160
DB 582 GATGTCTCGAGGCGCTTGTCTTTTTCACATGAGGCTATGTCATGACAGACCTCAAG 641
QY 161 ProArgAsnIleLeuTrpSerArgIleGluAsnGluCysPheLeuLeuAspPheGlyLeu 180
DB 642 CCACGAAACATCTCTGTGAGTGGCGAGAACGATGCTTTAAGCTTTATGACTTGGACATC 701
QY 181 SerPheLeuGlnGluYHisGlnAspValTYrTYrIleGlnThrAspGlyTYrArgAlaPro 200
DB 702 AGCTTCAAGAGAGCATCAGACGTAAGTATATTCAGACAGACGAGTATCAGACTCTCT 761
QY 201 GluAlaGlnLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerSerPheThrGluCysThr 220
DB 762 GAAGCAGAACTGAGAAATGCTTGGCCCGGCTGACAGTATACAGAGTGTACC 821
QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnIleuSerPheSerGlyMetCys 240
DB 822 TACGCTGTTCATCTTCGAGCCTTGCAGATCATTTTACTGGAATGTCTTCAGAAAGAA 881
QY 241 LeuLeuHisIleThrValArgSerGlnGluTYrPLeuValAsnSerSerAlaIleIleAspHis 260
DB 882 CTGAAACATATCAGATCATGATCTCAGAGTGAAGGCAACAGTTCTGCTATATTATGATCAT 941
QY 261 IlePheAlaSerIleValValAlaAlaIleProAlaTYrHisLeuAlaAspLeu 280
```

```
DB 942 ATATTGGCAGTAAGACGTGGTGAATGCCGCAATTCAGCTTACCTCAGACACTT 1001
QY 281 IleYsSerMetLeuHisAspAspProSerArgAlaIleProAlaGluMetAlaLeuCys 300
DB 1002 ATCAAAAGCATCTTCATGACACACCAAGCAGAAAGATCCCTGCTGAGATGGCTTGTGC 1061
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB 1062 AGCCCATCTTTAGCATTCCTTTGCCCCCTCATATTGAAGATCTGTGATGCTTCAACT 1121
QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTYrLeuGlyAsnGluGlnGlyTYr 340
DB 1122 CCAGTGCTAGACTCTCTCATATGCTGAGTGAATGACTATCTTGAAGATGAGATGAATAT 1181
QY 341 GluAspValAlaGluAspValTYrGlnGluCysGlnTYrGlyProValAlaSerLeu 360
DB 1182 GAAAGATGTTGTGAAGATGTAAAGAGTGTCAAGAAATATGACCAAGTGGTTTCTCTG 1241
QY 361 LeuValProLYsGluAsnProGlyArgGlyGlnValPheValGluTYrAlaAsnAlaGly 380
DB 1242 CTGTGTTCCAAAGAAATCTTGGCAGAGGACAAAGTCTTGTGAGTACGCAAACTGTGA 1301
QY 381 AspSerIleValAlaGlnLeuLeuThrArgIleArgMetPheAspGlyIlePheValVal 400
DB 1302 GATTCAAAGCTGTCTCAGAAATGCTGACTGGAGAGATGTTGACGGAGATTTGTGTG 1361
QY 401 AlaThrPheTYrProLeuSerAlaTYrIleYsArgGlyTYrLeuTYrGlnThrLeuLeu 419
DB 1362 GCTACATTTTACCGCTGAGTGTCTTACAAAGAGGAGATATCTTTATCAAACTGTGCTT 1418
RESULT 11
ADQ24401
ID ADQ24401 standard; DNA; 3998 BP.
XX
AC ADQ24401;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
PN MO2004048938-A2.
XX
ID 10-JUN-2004.
XX
PF 26-NOV-2003; 2003MO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnick A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 7221; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
```


CC cytostatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC DNA of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.

XX Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;

Alignment Scores:

| Pred. No.: | 2,296-192 | Length: | 3998 |
|------------------------|-----------|---------------|------|
| Score: | 2045.50 | Matches: | 396 |
| Percent Similarity: | 94.76% | Conservative: | 2 |
| Best Local Similarity: | 94.29% | Mismatches: | 21 |
| Query Match: | 92.72% | Indels: | 1 |
| DB: | 12 | Gaps: | 1 |

US-10-798-532-2 (1-419) x ADQ24401 (1-3998)

```

QY 1 MetAlaGlySerGlyCysAlaTTPGlyAlaGluProProArghPheLeuGluAlaPheGly 20
DB 160 ATGGCCGAGATCCGGCTGCGCTGGGGCGGAGCCCGCGTTCCTGAGGCGCTTCGGG 219
QY 21 ArgLeuTTPGlnAlaGlnSerArGLeuGlySerGlySerSerAlaSerValTYrArGVal 40
DB 220 CGGCTGTGGAGGATACAGAGCCCTGTGGTACGGGCTCCCTCCCGCTCGGTATCCGGTT 279
QY 41 ArgCysArgGlyAanProGlySerProProGlyAlaLeuGlyGlnPheLeuProProGly 60
DB 280 CGCTGTGGCGGCAACCTGGCTGCGCCCGCGGCGCTCAAGAGATTCCTGGCGGAGGA 339
QY 61 ThrThGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaAlaLeu 80
DB 340 ACCACCGGGGCTCGGCTCTGCGCGAGATGATTAAAGATGTTCTTACCAAGACG 399
QY 81 GluGlnLeuGlnGlyHisArgAanIleValThrLeuTYrGlyValPheThrIleHis--- 99
DB 400 AAGTCTGTGTTACCTTACCGAGAGTCACTTTGATGAGATGTTTCAATCCAGGTT 459
QY 100 PheSerProAanValProSerArGyCysLeuLeuGlnLeuLeuAanPValSerValSer 119
DB 460 GTTCTCAAGTGTGCATACCGCTGTCTGTCTTGAATCTCGATGTCAGTGTTCG 519
QY 120 GluLeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTrMetIleGlnHisCysAla 139
DB 520 GAATTCCTTAATATTCACATCAACGAGGTGTTCATGAGATGATGACGATTCGGC 579
QY 140 ArgAspValLeuGluAlaLeuAlaPheLeuHisIleGluGlyTYrValHisAlaAspLeu 159
DB 580 CGAGATGTTTGGAGGCGCTTGTCTTCTTCATCATGAGGCTATGTCCATGGGACCTC 639
QY 160 LysProArgAanIleLeuTYrSerArGlyAanGluCysPheLeuLeuIleAanPheGly 179
DB 640 AAACCAAGTAAATATGAGAGTGCAGAGAAATGAATGTTTAACTCATTTGACTTTGGA 699
QY 180 LeuSerPheLeuGluGlyAanGlnAanPValLysTYrIleGlnThrArgGlyTYrArGAla 199
DB 700 CTTAGCTTCAAAAGAGCAATCAGAGATGAATTAATTCAGACAGCGGTATCGGCT 759
QY 200 ProGluAlaGluLeuGlnAanCysLeuAlaGlnAlaGlyLeuGlnSerArPThrGluCys 219
DB 760 CCAGAGGAGCAATTCGAAATTCCTTGCGCCAGGCTGCTGCAGATGATACAGATGT 819
QY 220 ThrSerAlaValaAspLeuTYrSerLeuGlyIleIleLeuLeuGlnLecPheSerGlyMet 239
DB 820 ACCTCAGCTGTGATCTGTGAGACCTTGAATCATTTTACTGAAATGTTCTCAGGAATG 879
QY 240 LysLeuLysHisThrValaArgSerGlnGluTTPlyValaAnSerSerAlaIleIleAanP 259
DB 880 AAATGAAACATACAGTCAGATCTCAGGAATGAAAGCAAGTTCTGATTAATATGAT 939
QY 260 HisIlePheAlaSerLysAlaValaValaAlaIleProAlaTYrHisLeuArgASP 279

```

```

DB 940 CACATATTTGCCAGTAAAGCAATGGTGAATGCCGCAATTCACGCTTACCTTAAGAC 999
QY 280 LeuIleLysSerMetLeuHisArgAanProSerArGATGlieProAlaGluMetAlaLeu 299
DB 1000 CTATATCAAAAGATCTTCATATATATCCAGACAGAAATTCCTGCTGAATGGCATTG 1059
QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAanPLeuValMetLeuPro 319
DB 1060 TGCAGCCCATTTTATAGCATTCCTTTGCGCTCATATTTGAAGATCTGGTATGCTCC 1119
QY 320 ThrProValLeuArgLeuLeuAanValLeuAanPArPArTYrLeuGlyAanGluGlu 339
DB 1120 ACTCCAGTCTTAAGATCTGTAATGTGCTGATGATGATTTACTTGAGATGAAAGGAA 1179
QY 340 TYrGluAanPValValGluAanPValLysGluGluCysGlnLysTYrGlyProValAlaSer 359
DB 1180 TATGAAGATGTGTGAAGATTAAGAGGAGTCTCAAAATATATGACCACTGTATCT 1239
QY 360 LeuLeuValProLysGluAanProGlyArGlyGlnValPheValGluTYrAlaAanAla 379
DB 1240 CTACTTGTTCGAAAGAAATCTTGGCAGAGACAACTTTGTTGAGATGAAATGCT 1299
QY 380 GlysPserLysAlaAlaGlnLysLeuLeuThrGlyArGmetPheAspGlyLysPheVal 399
DB 1300 GGTGATTCGAAAGTCCGAGAAATTAATGACTGGAAGGATGTTGATGGGAAGTTTGT 1359
QY 400 ValAlaThrPheTYrProLeuSerAlaTYrLysArgGlyTYrLeuTYrGlnThrLeuLeu 419
DB 1360 GTGGCTACATTTCTACCCGCTGAGTGCATCAAGAGGAGTATCTGTATCAACCTTGT 1419

```

RESULT 12
 AAH70764
 ID AAH70764 standard; cDNA; 529 BP.
 AC AAH70764;
 XX
 DT 19-SEP-2001 (first entry)
 XX
 DE Human cervical cancer marker nucleic acid 2038.
 XX
 KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200142467-A2.
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000WC-US033312.
 XX
 PR 08-DEC-1999; 99US-0169681P.
 PR 21-DEC-1999; 99US-0171350P.
 PR 14-MAR-2000; 2000US-0189315P.
 PR 12-MAY-2000; 2000US-0203791P.
 PR 09-JUN-2000; 2000US-0210600P.
 PR 21-JUL-2000; 2000US-0220114P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Deeds J, Berger A, Zhao X;
 XX
 DR WPI, 2001-375006/39.
 XX
 PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 XX for assessing and detecting compounds for treating the cancer.
 XX
 PS Claim 1; Page 434; 1051p; English.
 XX
 CC The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;

CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

SQ Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 8,456-69 | Length: | 529 |
| Score: | 789.00 | Matches: | 159 |
| Percent Similarity: | 95.78% | Conservative: | 0 |
| Best Local Similarity: | 95.78% | Mismatches: | 5 |
| Query Match: | 35.77% | Indels: | 2 |
| DB: | 4 | Gaps: | 0 |

US-10-798-532-2 (1-419) x AAH70764 (1-529)

```
QY 220 ThSerAlaValAspLeuTPSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
DB 26 ACCTCAGCTGTTGATCTGGAGCCCTAGGAATCATTTTACTGGAAATGTTCTCAGGAATG 85
QY 240 LyeLeuIleHisThrValArgSerGlnIleTrpIleValAsnSerSerAlaIleIleAsp 259
DB 86 AAACCTGAACATACAGTCAGATCTCAGGAATGGAAGCAACAGTTCTGCTATTATTGAT 145
QY 260 HisIlePheAlaSerIleValAlaValAsnAlaAlaIleProAlaTrpHisLeuArgAsp 279
DB 146 CACATATTTTCCAGTAAGCAGTGGATGCGCAATTCAGCTTATCCTTACCTTAAGAGAC 205
QY 280 LeuIleIleSerMetLeuHisAspAspProSerArgIleIleProAlaGluMetAlaLeu 299
DB 206 CTATATCAAAAGCATGCTTCATGATGATCCAAAGAGAAATTCCTGCTGAATGGCAATG 265
QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
DB 266 TCGAGCCCATCTTTAGCATCTCTTTGGCCCTCATATTGAAATCTGTCATGCTTCC 325
QY 320 ThrProValIleuArgLeuAsnValIleuAspAspPtyrLeuGlyAsnGluGlu 339
DB 326 ACTCCAGTGTAGAGCTGAGTGAATGCTGATGATGATATCTTGAATGAAAGAGAA 385
QY 340 TyrGluAsp-ValValGluAspValIleuGlu-GluCysGlnIleTyrGlyProValValS 359
DB 386 TATGAAGATTGTTGTGAAGATGTAAAGAGAGAGTGTCAAAATATGAGCCAGGGGTAT 445
QY 359 LeuLeuValIleProGlyAsnProGlyArgGlyGlnValPheValGluTyrAlaAsn 379
DB 446 CTCTACTTGTCCAAAGGAAATCTCGCAGAGAAACAGTCTTTGTTAGTATGCAAAAG 505
QY 379 IagIyaAspSerIys 383
DB 506 CTGGGATTTCAAAG 519
```

RESULT 13

AAA02535
ID AAA02535 standard; cDNA, 722 BP.

AC AAA02535;

XX 19-MAY-2000 (first entry)

XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.

XX Human; colon cancer; tumour; diagnosis; gene expression product; probe;

XX detection; cancerous state; metastasis; identification; breast cancer;

XX oestrogen receptor-positive breast cancer; therapy;

XX oestrogen receptor-negative breast cancer; lung cancer; ss.

XX Homo sapiens.

XX WO958675-A2.

XX 18-NOV-1999.

PF 13-MAY-1999; 99WO-US010602.
XX
PR 14-MAY-1998; 98US-0085426P.
PR 15-MAY-1998; 98US-0085537P.
PR 15-MAY-1998; 98US-0085566P.
PR 21-OCT-1998; 98US-0105234P.
PR 27-OCT-1998; 98US-0105877P.

XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J,
PI Reinhard C, Gleese K, Randazzo F, Kennedy GC, Pot D, Kassam A,
PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I,
PI Leskowitz D, Kita D, Garcia V, Jones LW, Stache-Grain B;
XX MPI, 2000-126369/11.

XX Polynucleotide library used to determine cancerous states of mammalian
PT cells.

PS Claim 1; Page 1020; 1097pp; English.

XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived. The
CC polynucleotide sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of pre-
CC metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer

XX SQ Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 2,296-66 | Length: | 722 |
| Score: | 766.50 | Matches: | 182 |
| Percent Similarity: | 84.86% | Conservative: | 3 |
| Best Local Similarity: | 83.49% | Mismatches: | 20 |
| Query Match: | 34.75% | Indels: | 14 |
| DB: | 3 | Gaps: | 1 |

US-10-798-532-2 (1-419) x AAA02535 (1-722)

```
QY 152 GluGlyTyrValHisAlaAspLeuIleProArgAsnIleLeuTrpSerAlaGluAsn 171
DB 1 GAGGGCTATGTCATGCGGCTCAAAACCACTAATATTTGAGAGTGCAGAAATGAA 60
QY 172 CysPheIleLeuIleAspPheGlyLeuSerPheIysGluGlyAsnGlnAspValIleTyr 191
DB 61 TGTTTTAACTATGACTTGTGACTTANCTTCAAGAAAGCAATCAGGATTTAAAGTAT 120
QY 192 IlegInThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAla 211
DB 121 ATTCAAGACAGACGGGATCGGCTCCAGAAACA-GAATTGCAAAATTTGCTTGGCCANGCT 179
QY 212 GlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
DB 180 GGCCCTGCAGAGTGAACAAATGTAACCTCAGCTGTGATCTGTGAAGCTTACGAATCAT 239
QY 232 LeuLeuGluMetPheSerGlyMetIysLeuIleHisIleThrValArgSerGlnIleTrpIys 251
DB 240 TTACTGAAATGTTCTCAGGAATGAAACTGAACATACAGTCAAGATTCAGGAATGGAAG 299
```

QY 252 AlaAsnSer-SerAlaIleIle-AspHisIlePheHalaSerLysAla--ValValAsnAl 270
 Db 300 GCACACAGTTTCTGCTATATTTGATCACAATTTCCAGTAACCAANTGGTGAATGC 359
 QY 270 aaIaIleProAlaTyHisIleuArgAspLeuIleLysSerMetLeuHisAspAsp-ProS 290
 Db 360 CGCAATTCCAGCTTACACCTTANAGACCTTATCAAAAGAGATGCTTCATGATGCCAA 419
 QY 290 eArg-ArgIleProAlaGluMetAlaLeu-CysSerProPhePheSer-IleProPheA 309
 Db 420 GCAGGAAGATTTCTCTGTAATGGCATTTGTCACCCATCTTTAGCAATTCCTTTTG 479
 QY 309 lAProHisIleIleGluAspLeuValMetLeuProThrProVal-NeuArgLeuAsnVal 328
 Db 480 CCCCTATATTGAAGATCTGATCATCTTCCACCTCAGGCTAAGACTGCTGAATGTG 539
 QY 329 Leu-AspAspAspTyTyrLeuGlyValAsnGlu-GluGluTyTyrGluAspValValGluAspVal 347
 Db 540 CTGGGATGATGATTAATCTTGAGATGAAGAAGAGATTAAGAAGATGTTGTGAAGATGNT 599
 QY 348 LysGluGluCysGlnLys--TyrGlyPro 356
 Db 600 AAAAGAAGAGTGGCAAAATNTTGCACCA 629

RESULT 14

AAH71966
 ID AAH71966 standard; cDNA; 470 BP.

AC AAH71966;

DT 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 3240.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

OS Homo sapiens.

PN WO200142467-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000MO-US033312.

PR 08-DEC-1999; 99US-0169681P.

PR 21-DEC-1999; 99US-0171350P.

PR 14-MAR-2000; 2000US-0189315P.

PR 12-MAY-2000; 2000US-0203791P.

PR 09-JUN-2000; 2000US-0210600P.

PR 21-JUL-2000; 2000US-0220114P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Deeds J, Berger A, Zhao X;

DR WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and

PT for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 636; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with

XX cervical cancer with cytostatic activity. The nucleic acids and encoded

XX polypeptides are useful: to assess if a patient is afflicted with

XX cervical cancer or has a pre-malignant condition; to monitor the

XX progression of cervical cancer or a premalignant condition in a patient;

XX and to select and/or assess the efficacy of a compound or therapy for

XX inhibiting cervical cancer in a patient. The nucleic acids may also be

XX useful for gene therapy

XX Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6 51e-63 Length: 470
 Score: 729.00 Matches: 147
 Percent Similarity: 96.71% Conservative: 0
 Best Local Similarity: 96.71% Mismatches: 3
 Query Match: 33.05% Indels: 2
 DB: 4 Gaps: 0

US-10-798-532-2 (1-419) x AAH71966 (1-470)

QY 220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
 Db 18 ACCTCAGCGTGTGATCTGTGGAGCTTGAAGATCATTTTACTGGAATGTTCTCAGGAATG 77
 QY 240 LysLeuLysHisThrValArgSerGlnGluTyrPylAlaAsnSerSerAlaIleIleAsp 259
 Db 78 AAACGGAACAACACGTCAGATCTCAGAAATGGAAGCAACAGTTCGCTATTTATGAT 137
 QY 260 HisIlePheAlaSerLysAlaValValAlaAlaIleProAlaTyHisIleuArgAsp 279
 Db 138 CACATATTTGCCAGTGAAGCAAGTGTGAATGCCGCAATTCACCTTACCTTAAGAGAC 197
 QY 280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
 Db 198 CTTATCAAAAGCATGCTTCATGATGATCCAAAGAGAAATTCCTGCTGAATGCAATG 257
 QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
 Db 258 TGCAGCCCATTTCTTGAAGATTCCTTTGCCCCCTCATTTGAAGATCTGTCATGCTTCCC 317
 QY 320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyTyrLeuGlyValAsnGluGlu 339
 Db 318 ACTCCAGTCTTAAGACTGCTGAATGTGCTGATGATGATTAATCTTGAGAAAGAGAA 377
 QY 340 TyTyrGluAsp-ValValGluAspValLysGlu-GluCysGlnLysTyTyrGlyProValLys 359
 Db 378 TATGAAGATGTTGTGTAAGAAGATGTAAAGAGGAGTGTCAAAATATATGACCAAGGGTAT 437
 QY 359 eTyrLeuValProLysGluAsnProGlyArg 369
 Db 438 CTCTACTTGTTGCCAAAGAAATCCTGCGACA 469

RESULT 15

AAH69101
 ID AAH69101 standard; cDNA; 461 BP.

AC AAH69101;

DT 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 375.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

OS Homo sapiens.

PN WO200142467-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000MO-US033312.

PR 08-DEC-1999; 99US-0169681P.

PR 21-DEC-1999; 99US-0171350P.

PR 14-MAR-2000; 2000US-0189315P.

PR 12-MAY-2000; 2000US-0203791P.

PR 09-JUN-2000; 2000US-0210600P.

PR 21-JUL-2000; 2000US-0220114P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Deeds J, Berger A, Zhao X;

DR WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 168-169, 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

XX Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 1.e-53 | Length: | 461 |
|------------------------|--------|---------------|-----|
| Score: | 636.00 | Matches: | 138 |
| Percent Similarity: | 93.33% | Conservative: | 2 |
| Best Local Similarity: | 92.00% | Mismatches: | 5 |
| Query Match: | 28.83% | Indels: | 6 |
| DB: | 4 | Gaps: | 0 |

US-10-798-532-2 (1-419) x AAH69101 (1-461)

QY 220 ThrSerAlaValAspLeuTpsSerLeuGlyIleIleuLeuGluMetPheSerGlyMet 239
DB 22 ACCTCAGCTGTTATCTGTCGAGCCCTAGGAAATCATTTTACTGGAAAGTTCTCAGAAATG 81
QY 240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
DB 82 AAACCTGAACATACAGTCAGATCTCAGGAATGAGGCAAAACAGTTCTGCTATTATTGAT 141
QY 260 HisIlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTyrHisLeuArgAsp 279
DB 142 CACATATTGTCAGTAAGCAGTGTGAATGCCCAATTCAGCCTATCCCTAAGAGAGAC 201
QY 280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProIleGluMetAlaLeu 299
DB 202 CTTATCAAAAGCATGCTTCATGATGATCCAAAGCAAGAAATTCCTGCTGAATGGCATTG 261
QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
DB 262 TGCAGCCCATTTCTTAGCATTCCTTTGGCCCTCATAT- GAAGATCTGGTCATGCTTCC 320
QY 320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGluGlu 339
DB 321 ACTCCAGTGTCTAAGACTGCTGAATGCTGATGATGATTAATCTTGAGATGAAGAGAA 380
QY 340 TyrGluAsp--ValValGluAspValLysGlu--GluCysGlnLysTyrGlyPro--ValY 358
DB 381 TATGAAGATTGTGTTAAGAGATGAAGAGGAGAGGTGTCAAAAATATGACCAAGGTGG 440
QY 358 aLserLeuLeuValProLys 364
DB 441 TATCTCTACTTGTGTCCAAA 460

Search completed: November 29, 2004, 11:42:18
Job time : 521 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 3159.5 Seconds
(without alignments)
4832.489 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206
Sequence: 1 MAGSGCAGCABPPRFLEARG.....VATFVPLASVKKGYLTLL 419

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=xlp
-G/cgn2.1/USPTO.spool.p/US10798532/runat_29112004_112453_15901/app_query.faeta_1.1166
-DB=EST -QFMT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIT=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10798532_@CEN_1_1_6628_@runat_29112004_112453_15901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -XGAPOP=6
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gsa1.*
9: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 2174 | 98.5 | 1763 | 3 | AK013347 Mus muscu |
| 2 | 1711 | 77.6 | 2846 | 3 | AK030152 Mus muscu |
| 3 | 1521.5 | 69.0 | 1079 | 4 | BM468107 AGENCOURT |
| 4 | 1274.5 | 57.8 | 1015 | 6 | BY712344 BY712344 |
| 5 | 1242 | 56.3 | 917 | 4 | BM451184 AGENCOURT |
| 6 | 1170 | 53.0 | 823 | 7 | CO771206 test18_ES |
| 7 | 1162 | 52.7 | 779 | 5 | BU447831 BU447831 |
| 8 | 1159 | 52.5 | 896 | 7 | CO773091 test18 ES |
| 9 | 1133 | 51.4 | 781 | 7 | CN232453 CN232453 |

| | | | | | |
|----|--------|------|------|---|-------------------|
| 10 | 1110 | 50.3 | 932 | 5 | BU541306 |
| 11 | 1106 | 50.1 | 717 | 5 | BU85761 BU85761 |
| 12 | 1091 | 49.5 | 899 | 2 | BE895119 BE895119 |
| 13 | 1079 | 48.9 | 772 | 4 | BI62420 BI62420 |
| 14 | 1075 | 48.7 | 794 | 5 | BU205357 BU205357 |
| 15 | 1073.5 | 48.7 | 834 | 7 | CF593363 CF593363 |
| 16 | 1068.5 | 48.4 | 875 | 4 | BI259568 BI259568 |
| 17 | 1066 | 48.3 | 888 | 5 | BQ225071 BQ225071 |
| 18 | 1053 | 47.7 | 682 | 4 | BI333581 BI333581 |
| 19 | 1050 | 47.6 | 657 | 4 | BG547675 BG547675 |
| 20 | 1028 | 46.6 | 844 | 4 | BG474616 BG474616 |
| 21 | 1014 | 46.0 | 802 | 7 | CO737403 CO737403 |
| 22 | 1008.5 | 45.7 | 876 | 5 | BU184346 BU184346 |
| 23 | 990.5 | 44.9 | 906 | 4 | BG386905 BG386905 |
| 24 | 968.5 | 43.9 | 966 | 5 | BU439278 BU439278 |
| 25 | 967 | 43.8 | 631 | 5 | BU112717 BU112717 |
| 26 | 966.5 | 43.8 | 1003 | 5 | BQ222020 BQ222020 |
| 27 | 962 | 43.6 | 755 | 5 | BU371858 BU371858 |
| 28 | 958 | 43.4 | 699 | 5 | BU707284 BU707284 |
| 29 | 944 | 42.8 | 604 | 5 | BX261107 BX261107 |
| 30 | 909.5 | 41.2 | 716 | 7 | CF745774 CF745774 |
| 31 | 900 | 40.8 | 570 | 5 | BU293329 BU293329 |
| 32 | 896 | 40.6 | 670 | 7 | CN359673 CN359673 |
| 33 | 892 | 40.4 | 928 | 5 | BQ424438 BQ424438 |
| 34 | 880.5 | 39.9 | 824 | 2 | BE68528 BE68528 |
| 35 | 870.5 | 39.5 | 820 | 7 | CK597557 CK597557 |
| 36 | 862 | 39.1 | 672 | 5 | BX261106 BX261106 |
| 37 | 861.5 | 39.1 | 1063 | 6 | CD510246 CD510246 |
| 38 | 858 | 38.9 | 785 | 6 | CB990192 CB990192 |
| 39 | 850 | 38.5 | 523 | 2 | BF29410 BF29410 |
| 40 | 828.5 | 37.6 | 1063 | 5 | BU356643 BU356643 |
| 41 | 818.5 | 37.1 | 743 | 4 | BI562440 BI562440 |
| 42 | 818 | 37.1 | 763 | 7 | CK364431 CK364431 |
| 43 | 801 | 36.3 | 933 | 2 | BE886040 BE886040 |
| 44 | 796 | 36.1 | 986 | 5 | BU751496 BU751496 |
| 45 | 793 | 35.9 | 484 | 5 | BX281211 BX281211 |

ALIGNMENTS

RESULT 1
AK013347
LOCUS
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:281045404 product:kinase interacting with
leukemia-associated gene (etrachmin), full insert sequence.

ACCESSION AK013347
VERSION AK013347.1 GI:12850651
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P., and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kutsuna, T., Tachito, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,


```

|||||
Db 939 ATATTGCCAGTAAACAGAGTGGATGCGCAATTCAGCCTATCAGCTCAGAGACCTT 998
Qy 281 lIeIySeMeIeuthIsAspAspProserArArqIleProhIaIuMeIaIaIuCyS 300
Db 999 ATCAAAAGCATGTTCAATGATGACCCGGGAGAAAGATCCCTGCTAGATGGCATTTGTC 1058
Qy 301 SerProPheSerIleProPheAaProhIaIeGlUAspLeuValMeIeUpProThr 320
Db 1059 AGCCCATCTTTGCAATCTCTTTGCCCTCATATTGAAGATCTGGTGAATGCTCCGACT 1118
Qy 321 ProValIeuArgLeuLeuAenValIleuAspAspAspTyIleuGlYanngIuIuYr 340
Db 1119 CCACTGCTGAGGCTCCCTCAATGCTGCTGATGATGATTAATCTGAMAATGAAATGATATAT 1178
Qy 341 GlUAspValValGlUAspValIyGluIuIuCyGluIuYrIeYrIeYrProValIeSerIeu 360
Db 1179 GAAGATGTTGTAAAGATGGAAGAGAGAGTGTCAAGAAATATGACCAAGCTTTCTCTG 1238
Qy 361 IeuValProIyGluAspProGIYArGGLYGlValIeValIeIuYrIaIaAaIaIy 380
Db 1239 CTGTTCCAAAGAAATCTCTGGCAGAGGCAAGTCTTCTTGAATGCGGAACGCTGGT 1298
Qy 381 AspSerIySAlaIaIaIuIySLeuIeuthrGIYArGMePheAspGIYySPhaIVal 400
Db 1299 GATTCCAAAGCTGCTCAGAAATGTTGCTGACCTGGAGGATGTTTGAACGGGAAGTTGTTGTC 1358
Qy 401 AlEtThPheYrProIeuSerAlaIYrIeYrArGIYrIeYrIeYrIeYrIeYrIeYr 419
Db 1359 GCTACATTTACCCGCTGAGTGCCTACAGAGGGAGATCTTATCAAACTTGCTT 1415

RESULT 2
AK030152 2846 bp mRNA linear HTC 03-APR-2004
LOCUS DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4933401D07 product:kinase interacting with
leukemia-associated gene (statmin), full insert sequence.
AK030152
ACCESSION AK030152.1 GI:26326136
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10348636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Katsunoto, H., Sakaguchi, S., Ikegami, T., Kasaiwa, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the

```

```

PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12042159
AUTHORS Aichi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuta, M., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirozane, T.,
Hori, F., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohnato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

COMMENT
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 220-0045, Japan (E-mail:genome-res@gsr.riken.jp,
URL:htp://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:htp://genome.gsc.riken.jp/
URL:htp://fantom.gsc.riken.jp/.
Location/Qualifiers

FEATURES
source
1..2846
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4933401D07"
/db_xref="taxon:10090"
/clone="4933401D07"
/sex="male"
/cisue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
238..1230
/note="unnamed protein product; kinase interacting with
leukemia-associated gene (statmin) (MGI:1341908)
putative"
/codon_start=1
/protein_id="BAC26812.1"
/db_xref="GI:26326137"
/translation="MTLVFTTHSPVPSKCLLELDIVSVSELLYSHQGSM
MIOCARVLEALAFHHEGYHADLPENILMSKNECFKIDFGLSFKENQDVKY
IOTDGRAPAEALONCLAQIGSDTECSAVDLISGILLEMSSGMKLTYSOE
MKANSAPIDHIFASKAVVNAI PAYHEDLISKMIHDPGRIPAEMLCSPFSIP
FAPHEDLVMPLPTVIRILNVDDDDYLENEDEYEVDEPKCKYGVSVSLVKE
NGKQOVPEVYNAAGDSKRAQLTIGRMDEGFAVATFPLSAYRGVYQTL"
2623..2628
/note="putative"
2846
/note="putative"
/size="2846"

ORIGIN
polya_signal
polya_site
Alignment Scores:
Pred. No.: 2,78e-168 Length: 2846
Score: 1711.00 Matches: 326
Percent Similarity: 99.39% Conservative: 2

```

```

Beet Local Similarity: 98.79%  Mismatches: 2
Query Match: 77.56%  Indels: 0
DB: 3  Gaps: 0
US-10-798-532-2 (1-419) x AK030152 (1-2846)

QY 90 valThrLeuTygIValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
   :::
Db 238 ATGACTTATATGAGAGCTTATACATACACTTCTCTCCCATGTCGACATCAGCTGCTG
QY 110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrsSerHisGlnGly 129
Db 298 CTGCTTGAATCCTCGATGTCAGAGTGTCCGAATGCTTTTAAATTCCTACATCAGAGGC
QY 130 CysSerMetTrpMetIleGlnHisCysValArgAspValLeuGluAlaLeuAlaPheLeu 149
Db 358 TGCCTCATGTCGATGATACGACACTGCGCCGAGATGTTTGAGAGGCCCTTCCTTCTT
QY 150 HisHisGluGlyTyTyValHisAlaAspLeuLysProArgAsnIleLeuTPSerAlaGlu 169
Db 418 CACCATGAGGGCTACGTCATCGACAGACTCAAAACAGAAACATCCTGTGAGTGGAG
QY 170 AsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspVal 189
Db 478 AATGAGTGTTTAAGCTTATTTGACTTTGACTCAGCTTCAAGAAGGCATTCAGACGTA
QY 190 LysTyTrIleGlnThrAspGlyTyTyArgAlaProGluAlaGluLeuGlnAsnCysLeuAla 209
Db 538 AAGTATATTCAGACAGCGGATTCGGCTCCCTCGAGACAGACTCGAAGATCTTGCGC
QY 210 GlnAlaGlyLeuGlnSerAspThrGlnCysThrSerAlaValAspLeuTPSerLeuGly 229
Db 598 CAGCGCGGCTCGAGAGTGTATACAGAAATGTACCTCAGCTGTGATCTGTGAGACCTCGA
QY 230 IleIleLeuLeuGlnMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGlu 249
Db 658 ATCATTTTACTGAAAGTTCCTCAGAGATAGCTGAAACATACAGACATCTCAGGAA
QY 250 TrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAlaValAsn 269
Db 718 TGGAGGCAAAACAGTCTGCTATTTATTTGATCATATATTTCCAGTAAAGAGTGGTAAT
QY 270 AlaAlaIleProAlaTyTrHisLysArgAspLeuIleLysSerMetLeuHisAspAspPro 289
Db 778 GCGCGAATTCACACCTATCACCTCAGAGACCTTATCAAAAGCATGTGCTTCAATGACCG
QY 290 SerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAla 309
Db 838 GGCAGAAAGATCCCTGCTGATGAGTGTGAGCCCATTTCTTTAGCATTTCTTTTGCC
QY 310 ProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuAsnValLeu 329
Db 898 CCTCATATTAAGAAATCTGTGATGCTTCCAGCTCAGCTGCTCAGGCTCCCTCAATGTGCTG
QY 330 AspAspAspTyTrLeuGlyAsnGluGluTyTrGluAspValValGluAspValLysGlu 349
Db 958 GATGATGATTAACCTTGAATAATGAAGATGAATGAAGATGTTGAGAAAGATGGAAGAG
QY 350 GluCysGlnLysTyTrGlyProValValSerLeuLeuValProLysGluAsnProGlyArg 369
Db 1018 GAGTGTCTAGAAATATGAGCAGAGTGTCTGCTGCTTCCAAAGGAAATCCCTGGCAGA
QY 370 GlnGlnValPheValGluTyTyAlaAsnAlaGlyAspSerLysAlaAlaGlnLysLeuLeu 389
Db 1078 GGAACAAGTCTTCTGTGATGACGGAAGCTGTGTGATTCCAAAGCTGCTCAGAAAGTGTG
QY 390 ThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyTrProLeuSerAlaTyTr 409
Db 1138 ACTGGAGAGATGTTTGAAGGAGATTGTGTGGCTACATTTCACCGGTGAGTGGCTAC
QY 410 LysArgGlyTyTrLeuTyTrGlnThrLeuLeu 419
Db 1198 AAGAGGGAGATATCTTTATCAAAACCTTGCTT 1227

```

```

RESULT 3
BM468107
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12224 row: b column: 06
High quality sequence stop: 754.
FEATURES
Location/Qualifiers
1..1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5535653"
/risus_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
ORIGIN
Alignment Scores:
Pred. No.: 5,61e-149 Length: 1079
Score: 1521.50 Matches: 309
Percent Similarity: 92.86% Conservative: 3
Beet Local Similarity: 91.96% Mismatches: 17
Query Match: 68.97% Indels: 7
DB: 4 Gaps: 3
US-10-798-532-2 (1-419) x BM468107 (1-1079)

QY 90 valThrLeuTygIValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
   |||
Db 2 GTGACTTGTATGAGAGTGTTCATCACTTTCTCCAAATGTGCCATCAGCTGTCTG 61
QY 110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuTyrsSerHisGlnGly 129
Db 62 TTGCTTGAATCCTCGATGTCAGTGTGTTTGGAAATGCTTTATATTCAGTACCAAGGAT 121
QY 130 CysSerMetTrpMetIleGlnHisCysValArgAspValLeuGluAlaLeuAlaPheLeu 149
Db 122 TGTTCATGTCGATGATACGACATGTCGCCGAGAGTGTGGAGGCCCTTGCTTTTCTT 181
QY 150 HisHisGluGlyTyTyValHisAlaAspLeuLysProArgAsnIleLeuTPSerAlaGlu 169
Db 182 CATCATGAGGGCTATGTCATGCGACCTCAAAACAGTAACATATTTGAGTGGAG
QY 170 AsnGluCysPheLysLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGlnAspVal 189
Db 242 AATGAGTGTTTTAACTCATTTGATTTGACTTTGACTTCAAGAAGGCATCAGATGTA 301
QY 190 LysTyTrIleGlnThrAspGlyTyTyArgAlaProGluAlaGluLeuGlnAsnCysLeuAla 209

```



```

Db 302 AGTATATTGACGACACCGGTTATCGGGCTCCGAGAGCAGAAATTCGAAATTCCTGGCC 361
Qy 210 GlnAlaGlyLeuGlnSerAspThrGluCythrSerAlaValAspLeuTrpSerLeuGly 229
Db 362 CAGGCTGGGCTGAGAGTATAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
Qy 230 TlelleleuGlnuMetPheSerGlyMetLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 249
Db 422 ATCATTTTCTGGAATGTTCTGAGGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 481
Qy 250 TTPlyAlaAsnSerSerAlaIleleuAspHisIlePheAlaSerLeuValAlaValAsn 269
Db 482 TGGAAAGCAACAGCTTCTGCTATATTGATGATGATGATGATGATGATGATGATGATGAT 541
Qy 270 AlaAlaIleProAlaTyrrHisLeuArgAspLeuIleLeuSerMetLeuHisAspAspPro 289
Db 542 GCGCAATTCACAGCCATCATCCCTAAGAGACCTTATCAAAAGCATGCTTCATGATGATGATCA 601
Qy 290 SerArgArgIleProAlaGluMetAlaLeuCySerProPhePheSerIleProPheAla 309
Db 602 AGCAGAGAAATTCCTCTGTAATGGAATGTGTGAGCCCATCTTTTATGCAATCTCTTTGCC 661
Qy 310 ProHisIleGluAspLeuValMetLeuProThrProValIleuArgLeuLeuValIleu 329
Db 662 CCTCATATTGAAGATCTGGTCATGCTCCACATCCAGTGTGATGATGATGATGATGATGATG 721
Qy 330 AspAspAspTrpLeuGluAsnGluGluGluGluGluGluGluGluGluGluGluGluGlu 349
Db 722 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 781
Qy 350 GluCyGluGlyLeuTrpGlyProValIleSerLeuValProGluGluAsnProGlyArg 369
Db 782 GAGTGTCAAAATATGACACGAGGTATCTCTGATCTGATCTGATCTGATCTGATCTGATCTG 841
Qy 370 GlnGlnVal-PheValGluTrpAlaAsnAla-GlyAspSerLeuAlaAlaGlnLeu-Leu 388
Db 842 GGAACAAGCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901
Qy 389 LeuThrGlyArg--MetPheAspGlyLeuPheValAlaAla-ThrPheTyrrProLeu-- 406
Db 902 CTGACATCGAANAAGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 961
Qy 407 -SerAlaTyrrLeuArgGly---TyrrLeuTyrrGlnThrLeu 418
Db 962 GAGTGTCTTACAGAGAGGATTTATCTGATTCAAACCTT 1001

```

```

RESULT 4
BY12344 1015 bp mRNA linear EST 17-DEC-2002
LOCUS BY12344 RIKEN full-length enriched, 10, 11 days embryo whole body
DEFINITION Mus musculus cDNA clone 2810454004 5', mRNA sequence.
ACCESSION BY12344
VERSION BY12344.1 GI:27123621
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1015)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamakawa, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Balderelli, R., Hill, D.P., Bull, C.,
Hume, D.A., Quackenbush, J., Schramm, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V.,
Chochua, C., Corbani, L.E., Cousins, S., Dalla, R., Driscoll, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

```

Numata, K., Okido, T., Pavan, W.J., Perce, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kobayashi, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kogawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contract: Yoshinori Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-research.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN Integrated Sequence Analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

ORIGIN

Alignment Scores:

| | | | |
|------------------------|------------|---------------|------|
| Pred. No.: | 4, 99e-123 | Length: | 1015 |
| Score: | 1274.50 | Matches: | 252 |
| Percent Similarity: | 89.51% | Conservative: | 4 |
| Best Local Similarity: | 88.11% | Mismatches: | 23 |
| Query Match: | 57.77% | Indels: | 7 |
| DB: | 6 | Gaps: | 4 |

| QY | 339 | Glutylrctgluaspval1Val1GluaspVal1LysgluInguCysegluLysrYrYrGLyProvalVal1 | 358 |
|------------------------|---|---|-----------------------------|
| Db | 502 | GAATATGAAAGTGTGTGAGAAAGTAAAGAAAGAGAGTGTCAAAAATATGACCAAGTGTGA | 561 |
| QY | 359 | SerleuDeuVal1ProLysgLuaspProGly1YrgrglYgluVal1PheVal1GluYr1Alasn | 378 |
| Db | 562 | TCTCTACTGTTGTCCAAAGGAAATCTCGGACAGAGACAAAGCTTGTGTGAATGCAAAAT | 621 |
| QY | 379 | AlaGlyaspSer1Yr1Ala1Glu1YrLeuLeuThrGly1YrGmetPheaspGly1YrPhe | 398 |
| Db | 622 | GCTGGTATTCACAAAGCTGCCACAGAAATACTGACTGAGAGAGATGTTGATGGAGAGTTT | 681 |
| QY | 399 | Val1Val1AlaThrPheYr1ProLeuSerAla1YrLysArsg1Yr1YrLeuYrGlnThrLeu | 418 |
| Db | 682 | GTGTGGCTACATTCCTACCCGCTGAGTCCCTACAGAGAGGAGATATCTGATCAAACTTG | 741 |
| QY | 419 | Leu | 419 |
| Db | 742 | CTT | 744 |
| RESULT 6 | | | |
| LOCUS | CO771206 | | |
| DEFINITION | CO771206 | 823 bp | mRNA linear EST 04-AUG-2004 |
| ACCESSION | CO771206 | | |
| VERSION | CO771206.1 | GI:50972473 | |
| KEYWORDS | EST. | | |
| SOURCE | Gallus gallus (chicken) | | |
| ORGANISM | Gallus gallus | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus. 1 (bases 1 to 823) | | |
| AUTHORS | Shin, J. | | |
| TITLE | Testis cDNA library | | |
| JOURNAL | Unpublished (2004) | | |
| COMMENT | Contact: JiHye Shin Division of Animal Genetic Engineering School of Agricultural Biotechnology, Seoul National University San 56-1, Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea Tel: +82-31-479-7557 Fax: +82-31-479-7550 Email: jaecljojh@hotmail.net Seq primer: T3. Location/Qualifiers 1. .823 /organism="Gallus gallus" /mol_type="mRNA" /strain="Korean Native chicken" /db_xref="taxon:9031" /sex="Male" /tissue_type="Testis" /dev_stage="26weeks" /clone_lib="Testis cDNA library" /note="Vector: Uni-ZAP XR vector; Site_1: XhoI; Site_2: EcoRI; The libraries were synthesized using the ZAP(R)-cDNA synthesis method (Stratagene). cDNA was prepared, size-fractionated and inserted into the Uni-ZAP XR vector using XhoI linker-primer and EcoRI adaptor. After in vivo excision with E. coli strain SOLR, cDNA libraries from testis contain each insert size range of 0.5-3kb (n=20), and 0.5-3.5kb (n=20)." | | |
| FEATURES | | | |
| SOURCE | | | |
| ALIGNMENT SCORES: | | | |
| Pred. No.: | 3.61e-112 | Length: | 823 |
| Score: | 1170.00 | Matches: | 229 |
| Percent Similarity: | 84.72% | Conservative: | 15 |
| Best Local Similarity: | 79.51% | Mismatches: | 22 |
| Query Match: | 53.04% | Indels: | 22 |
| DB: | 7 | Gaps: | 2 |

US-10-798-532-2 (1-419) x CO071206 (1-823)
 Oy 47 GlySerProPro-----GlyAlaLeuLysGlnPheLeuProProGlyThrThrGly 63
 Db 6 GGAAGCTTCACCGCGCGTGGCGCGCGCTGTGAACCTAGTGAATCCCGGGCTGCAGAAAT 65
 Oy 64 ALaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeuGluGlnLeu 83
 Db 66 TCGGCACAGAGC----- 77
 Oy 84 GlnGlyAlaIslaArgAenlleValThrLeuTyrGlyValPheThrIleHisPheSerProAsn 103
 Db 78 -----GGCAACATCGTAGACTCTGTAGCGGCGTGTTCACCAACCACTACTCGGCCAAC 128
 Oy 104 ValProSerArgCysLeuLeuLeuGlnLeuLeuLeuAspValSerValSerGluLeuLeuLeu 123
 Db 129 GGGCCCGTCCCGCTCCCTGCTGCTGGAGCTCTGGATATTCAGCCGTTGTGAGCTGTCTGTG 188
 Oy 124 TyrSerSerHisGlnGlyCysSerMetTyrMetIleGlnHisCysValAspAspValLeu 143
 Db 189 CACTTCGACCAACCAAGGGCTGCTCATATGTGATATTCAGACTTCGCCCGCGACTCTGTG 248
 Oy 144 GluAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLysProArgAsn 163
 Db 249 GAAAGCCCTGGGCTTCTGTGACACCAAAAGGCTATCGTGCAGCAGCAAGCTGAAGCCACGCAAC 308
 Oy 164 IleLeuTyrSerAlaGluLeuGlnLysCysPheLysLeuLeuIleAspPheGlyLeuSerPheLys 183
 Db 309 ATCTGTGGAGCGCCGAGAGAGAGTGGTTTAAAGCTCATTAAGCTTTGAGCTTAAGCTTCAAA 368
 Oy 184 GluGlyValGlnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGlu 203
 Db 369 GAGGGGAAATCAGAGATGTGAATATATATTCAAACAGACGGGTATTCGGGCTCCAGAGGCAGAA 428
 Oy 204 LeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaVal 223
 Db 429 CTGCAGAAACTGCTTACAGACAGAGCGGGCTCCAGAGTGAAGAGAAATGACTCTGCTGTG 488
 Oy 224 AspLeuThrSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLysValLeuHis 243
 Db 489 GATCTGTGGAGTCTGGGAATGTTGTTTACTGAAATGTTCTCAGGAATGAACCTGAACAT 548
 Oy 244 ThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHisAlaPheAla 263
 Db 549 ACAATTCCAATTCAGGAATGGAAACAAACAGTTCTGCCATCATCATGCGATTTTTGGC 608
 Oy 264 SerLysAlaValValAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIleLysSer 283
 Db 609 AGTGAAGGGGGTGTAAATTCAGCCATTCAGAGCTTATCACTCAGAGACCTTATTTAAAGC 668
 Oy 284 MetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhe 303
 Db 669 ATGGCTTATGTGTGACCAAGGCAACCGAGCCCTCTGTCTAAAGAGCTTATATGAGGCCATTC 728
 Oy 304 PheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeu 323
 Db 729 TTCAGCAATTCCTTTGCTCTCCCAATTTGAAGATTGGTGAATGCTCCCAAGCCCTGTGCTG 788
 Oy 324 ArgLeuLeuAsnValLeuAspAsp 331
 Db 789 AGCGCTGCTGAATGTTCTAAACGAT 812
 RESULT 7
 BU447831 779 bp mRNA linear EST 29-NOV-2002
 LOCUS 60376757.F1 CSEQRBNI3 Gallus gallus cDNA clone CHESt695n17 5', mRNA
 DEFINITION sequence.
 ACCESSION BU447831
 VERSION BU447831.1 GI:25937142
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus

XR vector using XhoI linker-primer and EcoRI adaptor.
After in vivo excision with E. coli strain SOIR, cDNA
libraries from testis contain each insert size range of
0.5-3kb (n=20), and 0.5-3.5kb (n=20)."

ORIGIN

Alignment Scores:

| Pred. No.: | 5 89e-111 | Length: | 896 |
|------------------------|-----------|---------------|-----|
| Score: | 1159.00 | Matches: | 240 |
| Percent Similarity: | 81.90% | Conservative: | 18 |
| Best Local Similarity: | 76.19% | Mismatches: | 30 |
| Query Match: | 52.54% | Indels: | 27 |
| DB: | 7 | Gaps: | 2 |

US-10-798-532-2 (1-419) x CO773091 (1-896)

```

QY 46 ProglYser-PropPro-----GlyAlaLeuYsgInPheLeuProProGlyThrh 62
DB 11 CCTGGNACCTCCACCGCGTGGCGNCGCGCTTAAAGACTAGATGCCCGGGCTGCAG 70
QY 62 rglYAlaAlaAlaSerAlaAlaGluYrGlyPheArglyGlyUargAlaAlaLeuGlu 82
DB 71 GAATTCGGACGAGGC----- 86
QY 82 nLeuGInGlyHisArgAsnIleValThrLeuYrGlyValPheThrIleHisPheSerP 102
DB 87 -----CGGACATCGTGACTCTGTCGCGGCGTGTTCACCAACCACTACTCGGC 133
QY 102 oAenValProSerArgCyLeuLeuLeuGluLeuLeuAspValSerValSerGluLeu 122
DB 134 CAACGCGCGCTCCGCGCTGCTGCTGAGAGTGTGATATCAGCGTGTGCTGAGTGC 193
QY 122 uLeuYrSerSerHisGInGlyCySerMetTPrMetIleGInHisCyAlaArgAspVa 142
DB 194 GCGCACTCCAGCAACAGGCGCTGCTCAAGTGAATGATCCAGCACTGCGCCGACAC 253
QY 142 lLeuGluAlaLeuAlaPheLeuHisIleGluGlyYrValHisAlaAspLeuYsP 162
DB 254 GCTGGNAGCCCTGCTTCTCTGACACCAAGGCTACGTGCAGCACTGAGACCTGA 313
QY 162 gAenIleLeuTrpSerAlaGluAnGlyCyPheYsLeuIleAspPheGlyLeuSerPh 182
DB 314 CAACATCTGTGAGAGCGCGAGAGAGCTTTAAGCTCATGACCTTGGAATTGCTT 373
QY 182 elyGluGluYsArgIleAspValYsYrIleGInThrAspGlyYrArgAlaProGluAl 202
DB 374 CAAGAAGGGGAATCAGAGATGTGAATATATCAAAACAGCGGCTATCGGGCTC 433
QY 202 aGluLeuGInAsnCyLeuAlaGInIleGluGlnSerAspThrGluCyThrSerAl 222
DB 434 AGAATCTCGAATCTGCTTACACAGGCGGCTCCAGAGTGAAGAGGATGTACTCTGC 493
QY 222 aValaLeuLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetYsLeu 242
DB 494 TGTGATCTGTGAGCTGGGAATGTTTACTGGAATGTTCTCAAGGAATGAATGAA 553
QY 242 sHisThrValArgSerGInGluTrpYsAlaAsnSerSerAlaIleIleAspHisLe 262
DB 554 ACATTAAGTCCATCTCAGAAATGGAAGACAAACAGTTTCCATCATCATGCGCATTT 613
QY 262 eAlaSer-LysAlaValAlaAsnAlaAlaIleProAlaYrHisLeuArgAspLeuIle 282
DB 614 TGCAGTGAAGGGGGGTGTTAATTCAGCCATTCAGCTTATCCTCAGAGACCTTATTA 673
QY 282 ySerMetLeuHisAspAspProSerArgIleProAlaGluMetAlaLeuCySerP 302
DB 674 AAAGCATGCTTCACTGACCAAGCAAGGAGCGCTGCTGAAAGGCTTTATGAGCC 733
QY 302 roPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThPro 322
DB 734 CATTTCTTCAAGCATCTCTTGTCTCCCATTAAGAAAGTTTGATGAGCTCCACCGCTG 793
QY 322 alLeuArgLeu-LeuAsnValLeuAspAspTyrlLeuGlyAsn-GluGluGluYrG 341

```

```

DB 794 TCCTAGAGCTGACTGAATGTTTAAAGCATGCTTCTTCGACAGTGAAGAAATACGA 853
QY 341 uAspValValGlu-AspVallyGluGluCyGln 352
DB 854 AGAATCCCTGAGACATAGCGNAGAGAGTGTGAG 888

```

RESULT 9

LOCUS

CN232453 791 bp mRNA linear EST 09-APR-2004
WLB074D11.ab1 Wtestis Gallus gallus cDNA 5', mRNA sequence.

DEFINITION

CN232453

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EST analysis of brain and testis cDNA libraries from White Leghorn
and Red Jungle Fowl
Unpublished (2004)
Contact: Peter Savolainen
Department of Biotechnology
Royal Institute of Technology, KTH
SE-106 91 Stockholm, SWEDEN
Tel: +46 (0)8 5537 8481
Fax: +46 (0)8 5537 8335
Email: Peter.Savolainen@biotech.kth.se

Seq primer: M3 reverse primer.
Location/Qualifiers

FEATURES

source

1..781

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn"

/db_xref="taxon:9031"

/sex="male"

/lab_host="ElectroMAX DH10B (Invitrogen)"

/clone_lib="Wtestis"

/note="Organ: testis; Vector: pSPORT-1; Site 1: Hind III;

Site 2: EcoRI; The cDNA libraries were created with the

Superscript Plasmid System (Invitrogen)."

ORIGIN

Alignment Scores:

| Pred. No.: | 2.61e-108 | Length: | 781 |
|------------------------|--|---------------|-----|
| Score: | 1133.00 <td>Matches:</td> <td>225 </td> | Matches: | 225 |
| Percent Similarity: | 93.10% <td>Conservative:</td> <td>18 </td> | Conservative: | 18 |
| Best Local Similarity: | 86.21% <td>Mismatches:</td> <td>18 </td> | Mismatches: | 18 |
| Query Match: | 51.36% <td>Indels:</td> <td>3 </td> | Indels: | 3 |
| DB: | 7 <td>Gaps:</td> <td>0 </td> | Gaps: | 0 |

US-10-798-532-2 (1-419) x CN232453 (1-781)

```

QY 109 lLeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLeuYrSerSerHisGln 128
DB 2 CTGCTGCTGAGCTGCTGATATCAGCGTGTGAGCTGCTGCTCAGCTCCAGAACCG 61
QY 129 GlYCySerMetTrpMetIleGInHisCyAlaArgAspValLeuGluAlaLeuAlaPhe 148
DB 62 GCGCTGCTCATGTGATGATTCAGGACCTGCGCGCGAGCGTCTTGAAAGCCTTGCGCTTC 121
QY 149 LeuHisIleGluGlyYrValHisAlaAspLeuYsProArgAsnIleLeuTrpSerAla 168
DB 122 CTGCAACCAAAAGGCTACGTGACCAAGCCTGGAAGCAGCAATCTGTGAGCGGG 181
QY 169 GluAnGluCyPheYsLeuIleAspPheGlyLeuSerPheYsGluGlyAsnGlnAsp 188
DB 182 GAGGAGAGAGTGTCTTAAGCTCATGACTTGAATCAATCAAAAGAGGGGAATCAGAT 241

```

```
QY 189 ValIleTyrIleGlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCybLeu 208
Db 242 GTGAAATATATATCAACAGACGGGTATCGGGCTCCAGAGCAGAACTGCAGACTGGCTTA 301
QY 209 AlAGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTppSerLeu 228
Db 302 GCACAGGACGAGGCTCCAGAGTACGAGCAATGACCTGCTGATGATCTGTGAGAGCTG 361
QY 229 GlyIleLeuLeuGlnMetPheSerGlyMetLeuLeuYshIsthrValArgSerGln 248
Db 362 GGAATCTTTTACTGAAATGTTCTCAGAACTGAACCTGAACATCACTCAATCTCAG 421
QY 249 GluTyrPylalaAsnSerSerAlaIleAspHisIlePheAlaSerIleValVal 268
Db 422 GAATGAAAGCAAAACAGTTCTGCCATCATCATGATCTTTTCCAGTGAAGGGGTGTT 481
QY 269 AsnAlaAlaIleProAlaTyrHisIleuArgAspLeuIleYserMetLeuHisAspAsp 288
Db 482 AATTCAAGCCCTCAGTTATACCTCAGAGACCTTATTAAGCATGCTTCAATGTGAC 541
QY 289 ProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPhe 308
Db 542 CAGGCAAGACGACCTCTGCTGAAGAGCTTATGACGCCATCTTCAAGCATTCCTTT 601
QY 309 AlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuAsnVal 328
Db 602 GCTCCCATATTAAGATTTGGTGAATGCTCCACAGCTGCTGAGGCTGCAATGTT 661
QY 329 LeuAspAspAspIleuGlyAsnGluGluGluTyrGluAspValValGluAspValIle 348
Db 662 CTAGGCAATCTCTCTGCAAGAGTGAAGA-GAATACAAAGATATCTGGAAGA-ATTAAG 719
QY 349 GluGluCysGlnIleYsrGlyProValIleSerLeuLeuValProYsgIleAsnProGly 368
Db 720 GAGAGAGTCAAAATATGACCGGTGGTCTCTGCTTATTCGAGAG-ATTCCTGCT 778
QY 369 Arg 369
Db 779 AAA 781

RESULT 10
BUS41306 932 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
DEFINITION BUS41306
5', mRNA sequence.
ACCESSION BUS41306
VERSION BUS41306.1 GI:22851747
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 932)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9abrs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHC8765 row: 0 column: 18
High quality sequence stop: 631.
Location/Qualifiers
1..932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6572970"
```

```
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 40"
/note="Organ: prostate; Vector: pORF7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(6). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

ORIGIN

Alignment Scores:

| Pred. No.: | 8..94e-106 | Length: | 932 |
|------------------------|------------|---------------|-----|
| Score: | 1110.00 | Matches: | 228 |
| Percent Similarity: | 87.22% | Conservative: | 4 |
| Best Local Similarity: | 85.71% | Mismatches: | 18 |
| Query Match: | 50.32% | Indels: | 16 |
| DB: | 5 | Gaps: | 2 |

US-10-798-532-2 (1-419) x BUS41306 (1-932)

```
QY 1 MetAlaGlySerGlyCysAlaIleTrrGlyAlaGluProProArgPheLeuGluAlaPheGly 20
Db 149 ATGGCGGATCCGGCTGCGCGGAGCGCGCGCGGCTTCTGAGGCGCTTCGGG 208
QY 21 ArgLeuTrrGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db 209 CGGCTGTGGCAGGTACAGAGCCGTCTGCTGATACGGCTCTCCGCTCGGTATCGGTT 268
QY 41 ArgCysArgIleAsnProGlySerProProGlyAlaLeuLeuGlnPheLeuProGly 60
Db 269 CGGCTGTGGCAGAACCTGTGCTGCGCCCGCGCGCTTCAACATGTTCTTGGCCAGGA 328
QY 61 ThrThrGlyAlaAlaAsnSerAlaIleGluTyrGlyPheArgGlyGluAlaLeu 80
Db 329 ACCACCGGGGCTGCGGCTCTGCGCGGAGTATGTTCCGAAAGAGGGCGCGCTG 388
QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db 389 GAACAGTTCAGAGGTCAAGAAACATCGATCTTTGTATGAGAGTGTATCAATCCACTT 448
QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
Db 449 TCTCCAAATGTCCATCAAGCTGTGTGTTGTTGAATCTCTGAGATGTCAGTGTTCGGA 508
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrrMetIleGlnHisCysAlaArg 140
Db 509 TTGCTCTTATATTCAGTCACAGCGGTGTTCATGTGATATACAGCATTTGCGCGA 568
QY 141 AspValLeuGluAlaLeuAlaPheLeuHisIleGluGlyTyrValHisAlaAspLeu-Ly 160
Db 569 GATGTTTGGAGGCGCTTCTTCTCATCATGAGGGGTATGTCATTCGGAACCTCCGA 628
QY 160 sProArgAsnIleLeuTrrSerAlaGluAsnGluCysPheLeuLeuIleAspPheGlyLe 180
Db 629 ACCACGTAATATGTGAGTGCAGATGATGAATGTTTAAATCATGATGACTTTGGACT 688
QY 180 u-SerPheLeuGluGlyAsnGlnAspValIleGlnThrAspGlyTyrArgAlaP 200
Db 689 TTAGCTTCAAAAGAAAGCATCGAGTGAATATATTCAGCAAGCGGATTCGAGCTC 748
QY 200 roGlu-AlaGluLeuGlnAsnCybLeuAlaGlnAla-GlyLeuGlnSerAspThrGluCy 219
Db 749 CAGAAACAGAAATTCGAAATTTGCTTGCGCCAGCTGCGCAGAAAGATACAAATG 808
QY 219 sThrSer-AlaValAsp-LeuTrrSer-----LeuGlyIleIleLeuLeuGluMetPhe 236
Db 809 TACCTTCAGCTGTGATCTGTGAGAGCCCTAAGAAATCATTTTACCTGGAATAATGTCCTC 868
QY 237 SerGlyMetLeuLeuYshIsthrValArgSerGlnGluTrrPylalaAsnSerSerAla 256
Db 868
```

Db 869 CAGGATG-----GAAACTGGAAACTCAGCGTCCAGA 904

QY 257 llelle 258

Db 905 ATCTTC 910

RESULT 11

LOCUS BU385761

DEFINITION 717 bp mRNA linear EST 28-NOV-2002

ACCESSION 60382650F1 CSECHN75 Gallus gallus cDNA clone CHEST533d13 5', mRNA

VERSION BU385761

KEYWORDS BU385761.1 GI:25893762

SOURCE EST.

ORGANISM Gallus gallus (chicken)

REFERENCE Boardman, P.E., Sanz-Eguero, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J. A Comprehensive Collection of Chicken CDNAS Curr. Biol. 12 (22), 1965-1969 (2002)

AUTHORS

TITLE

JOURNAL

MEDLINE 22335534

PUBMED 12445392

COMMENT Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1. 717

source

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="white Leghorn, H1sex"

/db_xref="taxon:9031"

/clone="CHEST533d13"

/dev_stage="36"

/lab_host="DH10B"

/clone_lib="CSECHN75"

/note="Organ trunks; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores:

Pred. No.: 1.59e-105 length: 717

Score: 1106.00 Matches: 210

Percent Similarity: 93.72% Conservative: 14

Best Local Similarity: 87.87% Mismatches: 14

Query Match: 50.14% Indels: 1

DB: 5 Gaps: 0

US-10-798-532-2 (1-419) x BU385761 (1-717)

QY 134 Met15Gln1HisCys1AatgAapVal1LeuG1A1A1eua1A1a1eueu1His1G1u1G1y 153

Db 3 ATATATCAGACACTGCGCGCGGAGACGTCCTGGAAGCCCTTGCCCTGACCAACAAGGC 62

QY 154 TyrVal1HisAlaAapLeu1yProArGAsn1leLeu1rPSeAlaG1uAsnG1uCySpHe 173

Db 63 TACGTGACGACGACCTTGAAGCAGCAACATCCCTGTGAGGCGCGAGAGAGCTTT 122

QY 174 LysLeu1leAapPheG1yLeuSerPhe1yG1uG1yAaG1uAaPVal1yTr1leG1n 193

Db 123 AAGCTCATTAAGCTTGTGACTTGAAGAGGGGGAATCAGAGATGTGAATATATTCCA 182

QY 194 ThrAapG1yTyrArG1A1A1ProG1uA1G1uLeuG1uAaG1yAaG1uA1G1A1yLeu 213

Db 183 ACAGAGCGGTATCGGCTCCACAGAGCAGAACTGCACTGACACAGCAGCGCTC 242

QY 214 GlnSerAapThrG1uCyThrSerAlaValAapLeu1rPSeLeuG1y1leLeuLeu 233

Db 243 CAGAGTACAGCAATGTRACCTCTCTGTGATCTGTGAGCTGTGGAATTTTACTG 302

QY 234 G1uMetPheSerG1yMet1yLeu1ySh1rThrValA1ySerG1uTr1yAlaAa 253

Db 303 GAAATGTTCTCAGAATGAATGAACATACAGTCCATCTCAGGAATGMAAGCAAAAC 362

QY 254 SerSerAla1le1leAapHis1lePheAlaSer1yAlaValA1aA1a1lePro 273

Db 363 ACTTGTCCATCATCATGATCGCATTTTTCAGATGAAGGGGTGTAAATTCACGCTTCCA 422

QY 274 AlaTyrHis1leuArG1AapLeu1le1ySerMetLeuHisAapAapProSerArG1le 293

Db 423 GCTTATACCTCAGAGACCTTATTAAGCATCTTCTTGTGACCAAGCAAGCAGCC 482

QY 294 ProAlaG1uMetAlaLeuCySerProPhePheSer1leProPheAlaProHis1leG1 313

Db 483 TCTGCTGAAAGAGCTTTATGACGACCATTTCTCACATTTCTTGTCTCCCATATTGA 542

QY 314 AapLeuValMetLeuProThrProValLeuArG1eAaPVal1leuAaPheAapPyr 333

Db 543 GATTGTGATGATCTCCACGCTGTGAGCTGTGAAATGTTCTAAGCATCTTCT 602

QY 334 LeuG1yAaG1uG1uG1yTrG1uAaPVal1yG1uAaPVal1yG1uG1yG1yG1n 353

Db 603 CTGACAGAGTGAAGAAATACAGAAATCTCTGGAAGATTAAGGAGAGTGTAGAAA 662

QY 354 TyrG1yProVal1ySerLeuLeuValPro1yG1uAaPProG1yArG1yG1n 371

Db 663 TATGACCGGTGTTCTTCTTGTATTCGAAAGAAACATCTGTAAAGGCCAA 717

RESULT 12

LOCUS BE895119 899 bp mRNA linear EST 20-OCT-2000

DEFINITION 601436041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5', mRNA sequence.

ACCESSION BE895119

VERSION BE895119.1 GI:10358193

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 899)

AUTHORS NIH-MGC http://mgi.mc.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cga@bbs-rcmail.nih.gov

Tissue Procurement: ATCC/DC/DTP

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov

Plate: ILMA9753 row: b column: 12

High quality sequence stop: 626.

Location/Qualifiers

```
source
1. .899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3920939"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ALIGNMENT SCORES:
Pred. No.: 8 45e-104 Length: 899
Score: 1091.00 Matches: 210
Percent Similarity: 99.53% Conservative: 0
Best Local Similarity: 99.53% Mismatches: 0
Query Match: 49.46% Indels: 1
DB: 2 Gaps: 0

US-10-798-532-2 (1-419) x BE895119 (1-899)

QY 113 LeuLeuAspValSerValSerGluLeuLeuTySerSerHisGlnGlyCysSerMet 132
DB 1 CTCCTGGATGTCAGTGTTCGGAATGCTCTTATATCCAGTACACGAGGTTGTCATG 60

QY 133 TrpMetIleGlnHisCysAlaArgAspValLeuGluAlaPheLeuHisGlu 152
DB 61 TGGATGATACAGCATGTGTGCCAGAGATGTTTGAGGCGCTTCCTTCTTCATCATGAG 120

QY 153 GlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGlyCys 172
DB 121 GGGTATGTCATCGGAGCCTCAACACGTAACATNTGTGAGTGCAGAGATGATGT 180

QY 173 PheLeuLeuIleAspPheGlyLeuSerPheLeuGluGlyAsnGluAspValLysTyrIle 192
DB 181 TTAAATCTCATGTGACTTGTGAGCTTGTCAAGAAAGGCAATCAGATGTAAAGTATAT 240

QY 193 GlnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnGlyAsnAlaGly 212
DB 241 CAGACAGACGGGTATCGGCTCCAGAGAGCAAAATGCAAAATGCTTGCGCCAGGCTGGC 300

QY 213 LeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeu 232
DB 301 CTCGACAGATGATACAGAAATGATACCTGAGTGTGTGAGAGCCTAGAGAAATCATTTTA 360

QY 233 LeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTyrPheAla 252
DB 361 CTGGAAATGTTCTCAGGAATGAAACGTAAACATACAGTCTCAGGAATGGAAGGCA 420

QY 253 AsnSerSerAlaIleIleAspHisIlePheAlaSerLysValAlaValAsnAlaIle 272
DB 421 AACAGTCTGCTCATTTATGATACATATTTGCCAGTAAGACAGTGGGAATGCCGCAAT 480

QY 273 ProAlaTyrHisLeuArgAspLeuLysLeuSerMetLeuHisAspAspProSerArgArg 292
DB 481 CCAAGCCTATCACTAAGAGACCTTATCAAAAGCATGTCTTATGATATCCAAAGAGAGA 540

QY 293 IleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaProHisIle 312
DB 541 ATTCCTGCTGCTAAATGGCATTTGTGAGCCCATTTCTTTAGCATTCCTTTGGCCCTCATATT 600

QY 313 GluAspLeuValMetLeu-ProThrProVal 322
DB 601 GAAAGATCTGATCATGCTTTCCCACTCCAGTT 631

RESULT 13
BI562420 772 bp mRNA linear EST 05-SEP-2001
LOCUS BI562420
DEFINITION 60326316F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
mRNA sequence.
```

```
ACCESSION BI562420
VERSION BI562420.1 GI:15449747
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH_MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 772)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rt@mail.nih.gov
Tissue Procurement: Miklos Palokvits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1755 row: P column: 03
High quality sequence stop: 727.
Location/Qualifiers
1. 772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5298674"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptPR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC library."

ALIGNMENT SCORES:
Pred. No.: 1.23e-102 Length: 772
Score: 1079.00 Matches: 212
Percent Similarity: 96.82% Conservative: 1
Best Local Similarity: 96.36% Mismatches: 5
Query Match: 48.91% Indels: 3
DB: 4 Gaps: 1

US-10-798-532-2 (1-419) x BI562420 (1-772)

QY 92 LeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeuLeuLeu 111
DB 108 TTGTATGAGGTGTTTCAATCCACTTTCTCCAAATGTCACACGCTGTGTGTGCTT 167

QY 112 GluLeuLeuAspValSerValSerGluLeuLeuTySerSerHisGlnGlyCysSer 131
DB 168 GAATCTCTGATGTCTAGTCTTTCCGGAANTGCTCTTATATCCAGTACCAAGGTTGTTCC 227

QY 132 MetTrpMetIleGlnHisCysAlaArgAspValLeuGluAlaLeuAlaPheLeuHisHis 151
DB 228 ATGTGATGATACAGCATTTGTGCCAGATGTTTGGAGAGCCCTTCTTCTTCATCAT 287

QY 152 GlyGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGlu 171
DB 288 GAGGGCTATGTCCATGCGGACCTCAACACGTAACATATTTGTGAGTCAAGAAATGA 347

QY 172 CysPheLeuLeuIleAspPheGlyLeuSerPheLysGluGlyAsnGluAspValLysTyr 191
DB 348 TGTTTAACTCATTTGACTTGTGACTTACCTTCAAGAAAGGCAATCAGGATGTAAAGTAT 407
```


QY 192 IleglnThrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCySeuAlaGlnAla 211
Db 408 ATTCAAGACAGACGGGATTCGGGCTCCAGAGCAGAAATTCGAAATTCCTGGCCGAGCT 467
QY 212 GlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
Db 468 GGCCTGACAGATGATACAGAAATGATACCTGAGCTGTGATCTGAGAGCTAGAGATCAT 527
QY 232 LeuLeuGlnMetThrSerGlyMetLeuValLeuHisThrValArgSerGlnGluTrpLys 251
Db 528 TTAAGTGAATGTTCTCCAGAAATGAACTGAAACATACATGATCTCCAGAAATGAG 587
QY 252 AlaAsnSerSerAlaIleIleAspHisIle-PheAlaSerGlyAlaValAlaAsnAlaI 271
Db 588 GCAAACAGTTCTGCTATTATGATACATTAATTTGCCAGTAAAGCAGTGTGATGCCGC 647
QY 271 alleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerAr 291
Db 648 AATTCAGCCTTACCTAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAGACAG 707
QY 291 gaAgIleProAlaGluMetAlaLeuCySerSerProPhePheSerIleProPheAlaPro 310
Db 708 AAGAAATTCCTGCTGAATGGCATTTGGCAG-CCATTTCTTAGCA---TCCTTGCCCT 761

RESULT 14
BU205357 794 bp mRNA linear EST 25-NOV-2002
LOCUS 603104569F1 CSBQCHN03 Gallus gallus cDNA clone CHEST41J18 5', mRNA
DEFINITION sequence.
ACCESSION BU205357
VERSION BU205357.1 GI:25370848
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 794)
Boardman, P.E., Sanz-Exquerro, J., Overton, I.M., Burt, D.W., Boesch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.,
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
PUBMED 12445392

REFERENCE
AUTHORS Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
source

1. 794
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST41J18"
/issue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_11b="CSBQCHN03"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was bluntended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Alignment Scores:

| Score: | 3.38e-102 | Length: | 794 |
|------------------------|-----------|---------------|-----|
| Percent Similarity: | 1075.00 | Matches: | 221 |
| Best Local Similarity: | 89.10% | Conservative: | 16 |
| Query Match: | 83.08% | Mismatches: | 27 |
| DB: | 48.73% | Indels: | 4 |
| | 5 | Gaps: | 0 |

US-10-798-532-2 (1-419) x BU205357 (1-794)

QY 82 GlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSer 101
Db 1 CAGCTCCGGGGGACCGCAACATCGTACTGTACGGCGGTTCACCAACCACTACTCG 60

QY 102 ProAsnValProSerArgCysLeuLeuGlnLeuLeuAspValSerValSerGluLeu 121
Db 61 GCCAACGGCCCGTCCCGCTGCTGCTGAGCTGTGGATATACGCTGTGAGCTG 120

QY 122 LeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAsp 141
Db 121 CTGCTGCACTCCAGCAACAGCTCTCTCATGTGATATATCCAGCATCGCCCGGAC 180

QY 142 ValLeuGlnValLeuValPheLeuHisHisGlnGlyTyrValHisAlaAspLeuValPro 161
Db 181 GTCCTGGAAGCCCTGCGCTTCCGACACCAAAAGGCTACGTGCACGACGCAAGCA 240

QY 162 ArgAsnIleLeuTrpSerAlaGluAsnGluCysPheValLeuIleAspPheGlyLeuSer 181
Db 241 CGCAACATCTGTGAGCGCCGAGAGAGCTGTTAACTCATTTGACTTAC 300

QY 182 Phe-LysGlnGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaProG 201
Db 301 TTCAAAAGGGGAAATCAAGATGTGAAATATATTTCAAACAGCGGTATCGGCTCCA 360

QY 201 ValGlnLeuGlnAsnCySeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSe 221
Db 361 GGCAGAACTGCAAGAAATGCTAGACAGCAGGAGGCTCCAGATGAGACGGAATGTA 420

QY 221 ValValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys 241
Db 421 TGCTGTGATCTGTGAGCTGTGGAATGTTTAACTGGAATGTTCTCAGAAATGA 480

QY 241 uLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleAspHisI 261
Db 481 GAAACATACGTCATCTCAAGAAATGAAACAAACAGTTTGCATATCATGATCGCAT 540

QY 261 ePheAlaSerLysAlaValAlaAsnAlaIleProAlaTyrHisLeuArgAspLeuI 281
Db 541 TTTTGCAGTGAAGGGGTGTATATCCAGCATTCAGCTTATCACTCAGAGACCTTAT 600

QY 281 eLysSerMetLeuHisAspAsp-ProSerArgArgIleProAlaGluMetAlaLeuCys 301
Db 601 TAAACCATGCTTCATATGACCAAGCAAAAGCAAGCTCTGTGAAAAGCTTAATGCA 660

QY 301 ePhePhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThrP 321
Db 661 GCCCATTTCTCAGCATTCCTTGGCTCC-CATTTTAAATTTGGATGCTCCCAAG-C 718

QY 321 roValLeuArgLeuLeuAsnValLeuAspAspArgTyrLeuGlnGlnGlnGlnGlnGln 341
Db 719 CTGTGCTGAGGCTGTGATGTTCTTAAGCGCAAGGTTCTCGACAGATGAAGAAATACG 778

QY 341 LuAspValValGlu 345
Db 779 AGAATATCTGGGA 792

RESULT 15

CF593363
 LOCUS 834 bp mRNA linear EST 26-SEP-2003
 DEFINITION AGENCOURT 15622279 NIH_MGC.147 Homo sapiens cDNA clone
 IMAGE:30531268 5', mRNA sequence.
 ACCESSION CF593363
 VERSION CF593363.1 GI:36346826
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 834)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Garhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgarbs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 cDNA library preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAB621 row: m column: 05
 High quality sequence stop: 660.
 Location/Qualifiers
 1. 834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30531268"
 /tissue_type="Human Placenta"
 /lab_host="DH10B Tona"
 /clone_id="NIH_MGC_147"
 /note="Organ: placenta; Vector: pBluescriptR; Site 1:
 all-XhoI; Site 2: BamH; Oligo-dT primed using primer
 5'-TTTTTTTTTTTTTTVN-3', size-selected for average
 insert size 2.3 kb and normalized to R0T 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: This is
 a NIH_MGC library."

FEATURES

Source

ORIGIN

Alignment Scores:

Pred. No.: 5,24e-102 Length: 834
 Score: 1073.50 Matches: 214
 Percent Similarity: 95.13% Conservative: 1
 Best Local Similarity: 94.69% Mismatches: 5
 Query Match: 48.66% Indels: 6
 DB: 7 Gaps: 1

US-10-798-532-2 (1-419) x CF593363 (1-834)

QY 1 MetAlaGlySerGlyCyValATTPGlyAlaGluProBArgPheLeuGluAlaPheGly 20
 Db 155 ATGGCGGGAATCCGGCTCGCTGGGGCGCGAAGCGCGGCTTTCTGAGAGCCTTCGGG 214
 QY 21 ArgLeuTrpGluValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 215 CGGCTGTGGAGGACAGACCCGCTGTGGTGAAGCGGCTCCCGCTCGGTATCGGGTT 274
 QY 41 ArgCysGlyAlaSerProGlySerProProGlyValAlaLeuYsglnPheLeuProProGly 60
 Db 275 CGCTGCTGGCGCAACCTTGCTGCGCCCGCGGCCCTCAAGCAATTCTTGCGCGCAGGA 334
 QY 61 ThrThrGlyAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80

Db 335 ACCACCGGGGCTGCGGCTCTGCGCCGCGATGCTTCCGCAAGAGAGGGCGCGCTG 394
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 395 GAACAGTGGAGGTCACAGAAACATCGTACTTGTATGAGTGTATTACAACTCCACTTT 454
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuAspValSerValSerGlu 120
 Db 455 TTCCAAAAGTGCATCAGCGCTGTCTGTGCTTGAACCTCTGGATGTCAGTGTTCGGA 514
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysValAlaArg 140
 Db 515 TTGCTCTTATATTCAGTCACCGAGGTGTGTTCATGTGATGATACAGCAATTGTGCCGA 574
 QY 141 AspValLeuGlnValaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
 Db 575 GATGTTTGGAGGCTCTGCTTTCTTCATCATAGAGGCTATGTCCATCGAGCTCANA 634
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGly-Le 180
 Db 635 CCACTTAACATATTTGAGAGTGCAGAGAAATGAATGTTTAACTCATTTGACTTGGACTT 694
 QY 180 userPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArg-Alap 200
 Db 695 AGGCTTCAAGAAAGAACATCAGATGTAAAGTATTCAGACAGACGAGGTATCGGGGCTC 754
 QY 200 roGluAlaGluLeuGln-AsnCysLeuAla---GlnAlaGly-LeuGlnSerAspThrG 218
 Db 755 CAGAACAGAAATTCANAAATTCCTTGAGCCCAAGGCTTGCCCTTCAGAGTATACAGA 814
 QY 218 u-CysThr 220
 Db 815 AATGTACC 822

Search completed: November 29, 2004, 13:27:43
 Job time: 3172.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 26 Seconds
(without alignments)
1550.570 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206

Sequence: 1 MAGSCAMGAEPFRFILEARG.....VATFYPISAYKRGVLYQTLL 419

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------------|
| 1 | 1052 | 47.7 | 200 | 2 | I48615 gene KIS protein - |
| 2 | 248.5 | 11.3 | 948 | 2 | T24445 hypothetical prote |
| 3 | 243.5 | 11.0 | 720 | 2 | A96807 hypothetical prote |
| 4 | 243.5 | 11.0 | 826 | 2 | JC5153 mitogen-activated |
| 5 | 241.5 | 10.9 | 387 | 2 | T37758 protein kinase skp |
| 6 | 240 | 10.9 | 388 | 2 | S48879 protein kinase SMK |
| 7 | 235.5 | 10.7 | 367 | 2 | S68680 stress-activated p |
| 8 | 234.5 | 10.6 | 354 | 2 | T45138 protein kinase skp |
| 9 | 234.5 | 10.6 | 433 | 2 | S37790 probable serine/th |
| 10 | 234 | 10.6 | 435 | 2 | S64950 protein kinase HOG |
| 11 | 233 | 10.6 | 298 | 1 | A44878 protein kinase (EC |
| 12 | 233 | 10.6 | 298 | 2 | A41327 protein kinase (EC |
| 13 | 233 | 10.6 | 369 | 2 | A56492 protein kinase ERK |
| 14 | 230 | 10.4 | 373 | 2 | T13024 probable protein k |
| 15 | 230 | 10.4 | 376 | 2 | S40471 mitogen-activated |
| 16 | 230 | 10.4 | 427 | 2 | JC5693 stress-activated p |
| 17 | 228 | 10.3 | 423 | 2 | S43968 p54-alpha-2 stress |
| 18 | 227.5 | 10.3 | 469 | 1 | S17656 Ca2+/calmodulin-de |
| 19 | 227 | 10.3 | 326 | 2 | S23387 protein kinase (EC |
| 20 | 227 | 10.3 | 376 | 2 | S40470 mitogen-activated |
| 21 | 227 | 10.3 | 423 | 2 | S43967 p54-alpha stress-a |
| 22 | 227 | 10.3 | 424 | 2 | S71102 protein kinase JNK |
| 23 | 227 | 10.3 | 529 | 1 | S71774 calcium-dependent |
| 24 | 226.5 | 10.3 | 367 | 2 | JC5252 mitogen-activated |
| 25 | 226 | 10.2 | 346 | 1 | I48157 protein kinase (EC |
| 26 | 226 | 10.2 | 406 | 1 | KIHUCT phosphotyrase kina |
| 27 | 225.5 | 10.2 | 301 | 1 | KIHUCT protein kinase (EC |
| 28 | 225.5 | 10.2 | 371 | 2 | T14915 mitogen-activated |
| 29 | 225.5 | 10.2 | 416 | 2 | A48249 pre-mRNA splicing |

| | | | | | |
|----|-------|------|-----|---|---------------------------|
| 30 | 224.5 | 10.2 | 301 | 1 | A48041 protein kinase (EC |
| 31 | 224.5 | 10.2 | 678 | 2 | T43539 spindie checkpoint |
| 32 | 224 | 10.2 | 358 | 2 | S23383 protein kinase (EC |
| 33 | 223.5 | 10.1 | 335 | 2 | T23050 hypothetical prote |
| 34 | 223.5 | 10.1 | 342 | 2 | T21098 protein kinase (EC |
| 35 | 223.5 | 10.1 | 426 | 2 | S43968 p54-beta stress-ac |
| 36 | 223.5 | 10.1 | 464 | 2 | S71104 protein kinase JNK |
| 37 | 223.5 | 10.1 | 602 | 2 | S60052 calcium-dependent |
| 38 | 223 | 10.1 | 427 | 2 | JC5694 stress-activated p |
| 39 | 223 | 10.1 | 575 | 2 | JC7794 lammer kinase homo |
| 40 | 223 | 10.1 | 690 | 2 | T38052 probable protein k |
| 41 | 222.5 | 10.1 | 393 | 2 | S51321 mitogen-activated |
| 42 | 222.5 | 10.1 | 424 | 2 | A55480 c-Jun amino-termi |
| 43 | 222.5 | 10.1 | 473 | 1 | A53036 Ca2+/calmodulin-de |
| 44 | 222 | 10.1 | 346 | 1 | I78840 protein kinase (EC |
| 45 | 222 | 10.1 | 371 | 2 | S60121 mitogen-activated |

ALIGNMENTS

RESULT 1
I48615
gene KIS protein - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I48615
R/Maueuer, A.; Camonis, J.H.; Sobel, A.
Proc. Natl. Acad. Sci. U.S.A. 92, 3100-3104, 1995
A/Title: Stathmin interaction with a putative kinase and coiled-coil-forming protein dome
A/Reference number: I48282; MUID:95241452; PMID:7724523
A/Accession: I48615
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-200 <RES>
A/Cross-references: UNIPROT:P97343; EMBL:X82320; NID:g791078; PIDN:CAA57763.1; PID:g79107
A/Genetics:
A/Genes: KIS

Query Match 47.7%; Score 1052; DB 2; Length 200;
Best Local Similarity 99.0%; Pred. No. 1.3e-53;
Matches 198; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 122 | LYSSHQGSMMWICARDVLEALAPLHHEGYHADLKPRVILMSAENECFKLIDFGLS | 181 |
| DB | 1 | LYSSHQGSMMWICARDVLEALAPLHHEGYHADLKPRVILMSAENECFKLIDFGLS | 60 |
| QY | 182 | FKENQDVYKIOTDGYRAPEALONCLAQAGLSQDTECTSAVDLWSLGIILMFSGMKL | 241 |
| DB | 61 | FKENQDVYKIOTDGYRAPEALONCLAQAGLSQDTECTSAVDLWSLGIILMFSGMKL | 120 |
| QY | 242 | KHTVRSQEWKANSAAIIDHIFASKAVVNAATPAVHLRLIKSMHDDPSRRIPAEWALCS | 301 |
| DB | 121 | KHTVRSQEWKANSAAIIDHIFASKAVVNAATPAVHLRLIKSMHDDPSRRIPAEWALCS | 180 |
| QY | 302 | PFPSIPFAPHIEDVLMPLTP 321 | |
| DB | 181 | PFPSIPFAPHIEDVLMPLTP 200 | |

RESULT 2
T24445
hypothetical protein T04C10.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T24445
R/Burton, J.
submitted to the EMBL Data Library, March 1996
A/Reference number: Z19891
A/Accession: T24445
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-948 <WIL>

DB 332 --INDNTKYRVNRYRVVYFVIMRNKHFHSNVL-----NQGSKKEEK 372

RESULT 5

T37758 protein kinase ekp1p - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T37758

R:Jedler, H.; Wandurt, R.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, December 1998

A:Reference number: Z21744

A:Accession: T37758

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-387 <WEB>

A:Cross-references: UNIPROT:Q10452; EMBL:AL035064; PIDN:CAA22609.1; GSPDB:GN00066; SPDB:

A:Experimental source: strain 972h-; cosmid c1687

C:Genetics:

A:Gene: SPDB:SPAC1687.15

A:Map position: 1

A:Insertions: 7/1; 70/3

C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 10.9%; Score 241.5; DB 2; Length 387;
Best local Similarity 24.9%; Pred. No. 6e-07;
Matches 79; Conservative 47; Mismatches 112; Indels 79; Gaps 10;

QY 29 LSGSGSASVYRVRCNPGSPGALKQFLPPTGTAASAAYGFKERAALEFOLQGHN 88

DB 38 VSGSGFVVMQVHLISDSK--AAIKRVL-----QDKRFKRELQIWRIMKHEN 84

QY 89 IYLYGVF-----TIHSPNV---PSRCLLELDVSELLYSHOGS 131

DB 85 IYDLIYTYTGGNSEVYINLVLEFMPETTYRASRLYTRQKLSMPLLEVKL----- 137

QY 132 MNNIQCARDVLEALFLHHEGYVHADLKPRNIIIMSANECEFLIDFGLS--FKEGNODY 189

DB 138 -----IYQLRSLAVIHASGICHRDIKQNLIDPENQILKCPGSKILIVABEPNV 190

QY 190 KIIQTDGYRAPAEALONCLAQAGLQSDTECTSAVDLMSLGIILLMPFG----- 238

DB 191 SYICSRYYRAPEL-----IFGATDYTHAIDIMSTGCVAARMLGHPLPPGSGID 240

QY 239 -----MKLKTVASQEMKANSALIDHIFAS-----KAVNNAIPAYHLRDLIKSMUD 287

DB 241 QLVEIILKILGTPEERQIKTMNPNYMEHRFPQIRPPLSRVSRSPVLDAI--DLLSKLQY 299

QY 288 DPSRRIPAEALCSPFF 304

DB 300 TPTDRILTAEMKCHPFF 316

RESULT 6

S48879 protein kinase SMK1 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YP9499.10; protein YPR054w

C:Species: Saccharomyces cerevisiae

C>Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004

C:Accession: S48879; S54076

R:Striack, L.; Strich, R.; Winters, R.S.; Hall, J.P.; Mallory, M.J.; Keltzer, D.; Tuan,

Genes Dev. 8, 2151-2161, 1994

A:Title: SMK1, a developmentally regulated MAP kinase, is required for spore wall assem

A:Reference number: S48879; MUID:95047362; PMID:7958865

A:Accession: S48879

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-388 <KRI>

A:Cross-references: UNIPROT:P41808; EMBL:L35047; NID:G538516; PIDN:AAB59325.1; PID:G5385

R:Padcock, K.; Churcher, C.M.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54059

A:Accession: S54076

A:Molecule type: DNA
A:Residues: 1-388 <BAD>
A:Cross-references: EMBL:Z49219; NID:9805025; PIDN:CAA89172.1; PID:9805035; MIPS:YPR054w
A:Experimental source: strain AB972

C:Genetics:
A:Gene: SGD:SMK1

A:Cross-references: SGD:S0006258; MIPS:YPR054w

A:Map position: 16R

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphotransferase; signal transduction

F:36-306/Domain: protein kinase homology <KIN>

F:44-52/Region: protein kinase ATP-binding motif

Query Match 10.9%; Score 240; DB 2; Length 388;
Best local Similarity 25.8%; Pred. No. 7.3e-07;
Matches 103; Conservative 61; Mismatches 133; Indels 102; Gaps 19;

QY 21 RLMOVQSRIGSGSSASVYRVRCNPGSPGALKQFLPPTGTAASAAYG----- 72

DB 10 RAINVASNLGAPQQRITFAKERISIPGY--YEIIQFLKGAYGVCSVFKGRSPARIA 67

QY 73 -----PKKE---RAALEQL-----QCHNITVLYGVFTIHSPNVPSRCLLELDV 116

DB 68 VKKISNIFNKKEIILRAIRALEKFMNFKGKKNIVLIDLEIVTSPPDGLYC--YQELIDY 126

QY 117 SVSELLYSHOGSGMMNIQCARDVLEALFLHHEGYVHADLKPRNIIIMSANECEFLI 176

DB 127 DLAK-VIHSSVQ--LSEFHKITYQLGLCKTIHSADVIHRLKKNIGILCTL--NGCLKIC 183

QY 177 DFGLS-----FKEGNODYK-----YIOTDGYRAPAEALONCLAQAGLQSDTECTSAVDL 225

DB 184 DFLARGIHAGFPFKCHSTVQPHITVYVATRWYRADEL-----LISNQPYKSVDI 233

QY 226 NSGLGILILE-----MFSMKLKHIV-----RSQEW---KANSNAI 257

DB 234 MAVGCIILAEFYARKKVPFKWQRSDSMHQIFELIKVLGTPDDKILIKFTITAMNIGKSNNEV 293

QY 258 IDHI-----PASKAVNNAIPAYHLRDLIKSMUDPSRRIPAEALCSPFFSIFAP 310

DB 294 YKKIPMSNIFPPASHAIN-----LISLHMDSITHLNEQALSHFLNBYKRP 343

QY 311 HIEDVMLEPTPVRLINLVLDLDYLNBEYED--VVEDYK 348

DB 344 D-DEPVICLQGP-----FDFTYSELSMSKLRDYLVEYK 377

RESULT 7

S68680 stress-activated protein kinase (EC 2.7.1.-) 3 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S68680

R:Mertens, S.; Craxton, M.; Goedert, M.

FEBS Lett. 383, 273-276, 1996

A:Title: SAP kinase-3, a new member of the family of mammalian stress-activated protein j

A:Reference number: S68680; MUID:96198618; PMID:9925912

A:Accession: S68680

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-367 <MER>

A:Cross-references: UNIPROT:Q63518; EMBL:X96488; NID:G1262402; PIDN:CAA65342.1; PID:G1262

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphotransferase

F:25-280/Domain: protein kinase homology <KIN>

F:33-41/Region: protein kinase ATP-binding motif

Query Match 10.7%; Score 235.5; DB 2; Length 367;
Best local Similarity 22.8%; Pred. No. 1.2e-06;
Matches 96; Conservative 67; Mismatches 127; Indels 131; Gaps 19;

QY 10 APPRFLAEFGF-----LMQVQS-----RIGSGSSASVYRVRCNPGSPFG--ALK 54

DB 2 SSPPPAKGFYQGEVYTKTAMEVRAVYQDLQPVGSGAYGAV-----CSAVDSRTGKVAIK 56

QY 55 QFLPCTGAAAGAAEYGFKEKRALQLOGHNIYLYGVFTIHSPVNSRC----- 108
 DB 57 KLYRPFQSELFARKA-----YRELRLKHKM-HENVIGLDVFT-----PDETLDDFD 104
 QY 109 --LLELLDVSVSELLLYSHQCSMMIIOHCARDVLEALFLHNGYVHADLKRNILM 166
 DB 105 FYLVMPFMDGLCKLM---KHETLSBDRLQFLVYOMKGGIKYIHAAGVHIRDKPNL-- 159
 QY 167 SAENEC-FKLIDRGLSPKGNODVKYIOTDGYRPAPEALONCLAQAGLOSTECTSAVDL 225
 DB 160 ANVEDDELKLDGLARQADSEMTGYVTRWYPAPEYL-----NMKRYTQTVDI 209
 QY 226 WSLGILLLENFSGMKLKHTRVSGEMKANSALIDHIFASKAVYNA----- 270
 DB 210 MSVGCMAEMITGKIL-----FKGN-----DHLDQLEIKMKVGTPEPFVQKLSA 256
 QY 271 -----ALPAYLRD-----LISMLHDPSSRIIPAMALCSFPFSLPRA 309
 DB 257 EAKNMEGPELEBKDFASVLTNAPQAVNLKRMVLDEQHVTAALAHPEF-- 312
 QY 310 PHIEDLVMLPTVLRLLNVLDVDDYLGNEBEYEDVEDVK---EECOK--YGPVSLVLPK 364
 DB 313 -----SLKDTDEPKAQKIDSPDVRTLEEMKRVYKXVLSFRKPR 355
 QY 365 E 365
 DB 356 Q 356

RESULT 8

T45138
 protein kinase skp1 [imported] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T45138
 R:Polye, S.E.; Feoktistova, A.; Burke, J.D.; Woodgett, J.R.; Gould, K.L.
 Mol. Cell. Biol. 16, 179-191, 1996
 A>Title: Schizosaccharomyces pombe skp1+ encodes a protein kinase related to mammalian g
 A:Reference number: 222927; MID:96104567; PMID:8524294
 A:Accession: T45138
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-354 <P>Y>
 A:Cross-references: EMBL:L29449; NID:G1914882; PIDN:AAB51081.1; PID:G1914883
 C:Genetics:
 A:Gene: skp1
 A:Introns: 6/3
 C:Function:
 A:Pathway: cell's w
 A>Note: cells with deletions in skp1 are sensitive to heat shock and exhibit defects in
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: phosphoprotein; phosphotransferase
 F:353/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 10.6%; Score 234.5; DB 2; Length 354;
 Best Local Similarity 24.6%; Pred. No. 1.4e-06;
 Matches 78; Conservative 47; Mismatches 113; Indels 79; Gaps 10;

QY 23 LSGSSASVYRVCCNGPSPPALQFLPRTTGAAGAAEYGFKEKRALQLOGHNI 88
 DB 38 VSGSGFVWQVHLIESDK-AAIKRVL-----QDKRFKRELQIKRMINDPN 84
 QY 89 IYVLYGVF-----TIHSPV---PSRCLLELDVSVSELLLYSHQGS 131
 DB 85 IYVLIAYYTTGNSDEVYINLVLEFMPETIYRASRIYTKQKSMPLVYKLT----- 137
 QY 132 MMNIQHCARDVLEALFLHNGYVHADLKERNILMSAENECFKLIDFGLS--FKGNQDV 189
 DB 138 -----IYQLRSLAYIHASGICHRDIKPNLLDPENGILKLCDFGSAKILVAGBPV 190
 QY 190 KYIOTDGYRPAPEALONCLAQAGLOSTECTSAVDLMSLGIILLFMSG----- 238

DB 191 SYICSRYYRPAEL-----IFGADYTHAIDIWSTGCVMAELMCHLPFPEGSGTD 240
 QY 239 -----MKLKHTRVSGEMKANSALIDHIFAS-----KAVNAALPAYLRDLKSMLEHD 287
 DB 241 QLVETIKIKIGTSPREGIKTMNNVMEHRRPQLRPQLSVFSPRSVPLDAL-DLISKMLQY 299
 QY 288 DPSRRIPAMALCSFPF 304
 DB 300 TPTDRLTAAMAMCHPFF 316

RESULT 9

S37790
 probable serine/threonine-specific protein kinase (EC 2.7.1.-) YKL161c - yeast (Saccharom
 N:Alternate names: protein YKL615
 C:Species: Saccharomyces cerevisiae
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Aug-2004
 C:Accession: S37790; S37991; S44567
 R:Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.
 submitted to the EMBL data library, September 1993
 A>Description: DNA sequencing of a 36.2 kb fragment located between the PAS1 and LAP4 loc
 A:Reference number: S37786
 A:Accession: S37790
 A:Molecule type: DNA
 A:Residues: 1-433 <VAN>
 A:Cross-references: UNIPROT:P36005; EMBL:226877; NID:G407482; PIDN:CAA81493.1; PID:G40748
 A:Experimental source: strain S288C
 R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
 submitted to the Protein Sequence Database, March 1994
 A:Reference number: S37976
 A:Accession: S37991
 A:Molecule type: DNA
 A:Residues: 1-433 <VA2>
 A:Cross-references: EMBL:228161; NID:G486280; PIDN:CAA82003.1; PID:G486281; MIPS:YKL161c
 A:Experimental source: strain S288C
 R:Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.
 yeast 10, 35-40, 1994
 A>Title: DNA sequencing of a 36.2 kb fragment located between the PAS1 and LAP4 loci of c
 A:Reference number: S44563
 A:Accession: S44567
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-433 <VA3>
 A:Cross-references: EMBL:226877; NID:G407482; PIDN:CAA81493.1; PID:G407487
 A:Experimental source: strain S288C
 C:Genetics:
 A:Cross-references: SGD:S0001644
 A:Map position: 11L
 A>Note: YKL161c
 C:Superfamily: protein kinase homology
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F:21-287/Domain: protein kinase homology <KIN>
 F:153/Active site: Asp #status predicted

Query Match 10.6%; Score 234.5; DB 2; Length 433;
 Best Local Similarity 29.5%; Pred. No. 1.7e-06;
 Matches 89; Conservative 45; Mismatches 97; Indels 71; Gaps 16;

QY 75 KERALELOGHNIYLYGVFTIHSPVNSRC-----ELLDVSVSELLLYSHQCSM 132
 DB 71 RELKLRHLDHGNPNVWLDLT-DIVFYNGALNGYLYBELMECDISQIT--RSEQRLED 127
 QY 133 MMNIQHCARDVLEALFLHNGYVHADLKERNILMSAENEC-FKLIDFGLS-----FKEG 185
 DB 128 AHQSPRIYQILCAKRYIHSAVNLHODKRNKL--VNSDQCKICNFGLSGSENNKVN 185
 QY 186 NQDVK-YIOTDGYRPAPEALONCLAQAGLOSTECTSAVDLMSLGIILLF-----MBSG 238
 DB 186 DGIKGYIYISYKAPET-----LNLNOCETKAVIDISTGCIILAEELGRRKMPFG 235
 QY 239 M-----KLKHTR-----SGEMKANSALIDHIF-----ASKAVNAALPAYLR 277
 DB 236 KQYVDHLNIIQLIGTPPETLQE--IASQKYNIIIFQGNIPGSRFSFSLPGANP--EA 291

QY 278 RDLTKSMHDDPSRRIPAEWALCSPPFSI-----PFA-----PHIEDLWMLPTV 322
 Db 292 LELTKKMLERDPKRIITVEDALSHPIUSMWDIDEEFSCQITRFERFHEHESMAELGNEY 351
 QY 323 LR 324
 Db 352 IK 353

RESULT 10

S64950
 protein kinase HOG1 (EC 2.7.1.-), mitogen-activated - yeast (Saccharomyces cerevisiae)
 N/Alternate names: protein l2931; protein YLR113W
 C/Species: Saccharomyces cerevisiae
 C/Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 16-Aug-2004
 C/Accession: S64950; S42145; S69400
 R/Vername: P.; Voet, M.; Voelckert, G.
 Submitted to the Protein Sequence Database, May 1996
 A/Reference number: S64943
 A/Accession: S64950
 A/Molecule type: DNA
 A/Residues: 1-435 <VER>
 A/Cross-references: UNIPROT:P32485; EMBL:Z73285; NID:g1360507; PIDN:CAA97680.1; PID:e245
 A/Experimental source: strain S288C
 R/Brewer, J.L.; de Valoir, T.; Dwyer, N.D.; Winer, E.; Gustin, M.C.
 Science 259, 1760-1763, 1993
 A/Title: An osmoregulating signal transduction pathway in yeast.
 A/Reference number: S42145; MUID:93206121; PMID:7681220
 A/Accession: S42145
 A/Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 1-408 'GORGSCSK'

 A/Cross-references: EMBL:L06279; NID:g295610; PIDN:AAA34680.1; PID:g295611
 R/Vername: P.; Voelckert, G.
 submitted to the EMBL Data Library, September 1995
 A/Reference number: S69393
 A/Accession: S69400
 A/Molecule type: DNA
 A/Residues: 1-435 <VER>
 A/Cross-references: EMBL:X89514; NID:g1297019; PIDN:CAA61691.1; PID:e198746; PID:g129702
 C/Genetics:
 A/Genes: SGD:HOG1; SSK3
 A/Cross-references: SGD:S0004103; MIPS:YLR113W
 A/Map position: 12R
 C/Superfamily: protein kinase homology
 C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase; signal tr
 F;21-271/Domain: protein kinase homology <KIN>
 F;29-37/Region: protein kinase ATP-binding motif
 F;144/Active site: Asp #status predicted

Query Match 10.6%; Score 234; DB 2; Length 435;
 Best local similarity 25.7%; Pred. No. 1.8e-06;
 Matches 98; Conservative 53; Mismatches 140; Indels 90; Gaps 18;

QY 17 EAFGRIMQVQSR-----LGGSSASVYRRCGNNP--GSPPGALQFLPPTGTAAGA 68
 Db 11 QIFGVPELTNRNDINPVGMGAFGLV---CSATDTLTSPPAIKIKIMPSTAVLAKR 66
 QY 69 AEGYFKERPALEEQLOGRNIVTLVGVFTIHSPNVPSCRLLLELDVSVSELLYSHQ 128
 Db 67 T-----YRELKTLGHLR-HENTLICLDQIF---LSP-LBDIYFVTELGTDHLRL---QTR 114
 QY 129 GCSMMWIOICARVLEALAFIHHGIVYADLKPRNIIWSAENECFKLIDFGLSFKEGND 188
 Db 115 PLEKQVQYFLYXOILGKLKTVHSAVYHRDLKPSNIIIN-ENCDLXICDGLRIODPQW 173
 QY 189 VKYIOTDGYAPARAEIIONCLAQAGLSDPTCTSAVDLMSIGIILLEMFSGMK-----KHT 244
 Db 174 TGVVSTRYTRAPRIML-----TWQKIDV---VDIWSAGCI-PAEMIBSKPLPFGKDH- 222
 QY 245 VRSQEWKANSASAIIDHIFAS--RAVVAALPAVHLR----- 278

Db 223 -----VHQSFIITDILGSPKQVINTICSENTLKFTVSLPHRPDIPFSERFKTVEBPA 275
 QY 279 -DLTKSMHDDPSRRIPAEWALCSPPFSIPRPHIEDLV-----LPTPVRL 325
 Db 276 VDLTKMLVFPDKRIITADALAHF-YSAFYHPTDEPVADKFDPMHFNADLPVDITWKV 334
 QY 326 L---NVLDDVYLGNEEVEDV 343
 Db 335 MMYSRILDPHKIGSGDQIDI 355

RESULT 11

A44878
 protein kinase (EC 2.7.1.37) cdk2 [validated] - goldfish
 C/Species: Carassius auratus (goldfish)
 C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
 C/Accession: A44878
 R/Hirai, T.; Yamashita, M.; Yoshikuni, M.; Tokumoto, T.; Kajitara, H.; Sakai, N.; Nagahama
 Dev. Biol. 152, 113-120, 1992
 A/Title: Isolation and characterization of goldfish cdk2, a cognate variant of the cell c
 A/Reference number: A44878; MUID:92331802; PMID:1339336
 A/Accession: A44878
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-298 <HIR>
 A/Cross-references: UNIPROT:P43450; GB:S40289; NID:g251619; PIDN:AA822550.1; PID:g251620
 A/Experimental source: oocyte
 A/Note: sequence extracted from NCBI backbone (NCBIN:108782, NCBI:P108783)
 C/Genetics:
 A/Genes: cdk2
 C/Superfamily: kinase-related transforming protein; protein kinase homology
 C/Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine/
 F;2-25/Domain: protein kinase homology <KIN>
 F;10-18/Region: protein kinase ATP-binding motif
 F;33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 10.6%; Score 233; DB 1; Length 298;
 Best local similarity 28.7%; Pred. No. 1.4e-06;
 Matches 99; Conservative 48; Mismatches 116; Indels 82; Gaps 17;

QY 16 LEAFGRIMQVQSRILGSSASVYRRCGNNPSPGALQF-LPPTGTAAGAAYGRR 74
 Db 1 MSF-----QKVKIKIGEGTVGVVYKAK--NKVTGETVALKRIIDTREGVPTAL----- 49
 QY 75 KRALAEQLOGRNIVTLVGVFTIHSPNVPSCRLLLELDVSVSELLYSHQGSCMM 134
 Db 50 RSLSLKEL-NRPNTVYKLDV--IHTENK---YVFEFLHODLKRFRMDSVTGTISLP 103
 QY 135 IOHCARDVLEALAFIHHGIVYADLKPRNIIWSAENECFKLIDFGLSFKEGNDVYK--- 191
 Db 104 VKSYLFQDLQGLAFCHSHRVLRDLKPOYLINAGGE-IKLADFLAARAFGVPAVTTYHE 162
 QY 192 IOTDGYAPARAEIIONCLAQAGLSDPTCTSAVDLMSIGIILLEMFSGMKIKHTVRSQEWK 251
 Db 163 VVTLWYRAPEI-LIGC-----KYVSTAVDIMSIGIFAEWIT-----RKALPP 204
 QY 252 ANSALIDHIFPKSKAVN-----AATPAHL-----IDL 280
 Db 205 GSSE--IDQLFRIFPTLTGPDSEIWGVYSMDYKSPFKMARODLSKVPELDEGRDL 262
 QY 281 IKSMHDDPSRRIPAEWALCSPPFSIPFAPHTEDLVMLPTVRL 325
 Db 263 LQGMILYDPNKRIISAKNALVHRF-----RDVTM-FVPEPLRL 298

RESULT 12

A41227
 protein kinase (EC 2.7.1.37) cdk2 - human
 N/Alternate names: Egl homolog; protein kinase p34
 C/Species: Homo sapiens (man)
 C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
 C/Accession: A41227; S17873; S16520
 R/Ninomiya-Tsuji, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.

Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991

A:Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of a

A:Reference number: A41227; MUID:92020980; PMID:1717994

A:Accession: A41227

A:Molecule type: mRNA

A:Residues: 1-298 <NINP>

A:Cross-references: UNIPROT:P24941; GB:M68520; NID:g180177; PID:AAA5667.1; PID:g180178

R:Itai, L.H.; Harlow, E.; Meyerson, M.

Nature 353, 174-177, 1991

A:Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1A-

A:Reference number: S17873; MUID:91367262; PMID:1653904

A:Accession: S17873

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-298 <TSA>

A:Cross-references: GB:X62071; NID:g312802; PIDN:CAA43985.1; PID:g312803

R:Elledge, S.J.; Spottswood, M.R.

EMBO J. 10, 2653-2659, 1991

A:Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc28

A:Reference number: S16520; MUID:91330891; PMID:1714386

A:Accession: S16520

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-176, 'S', 178-298 <ELU>

A:Cross-references: EMBL:X61622; NID:g29848; PIDN:CAA43807.1; PID:g29849

C:Genetics:

A:Gene: GDB:CDK2

A:Cross-references: GDB:128984; OMIM:116953

A:Map position: 12q13-12q13

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine

F:2-255/Domain: protein kinase homology <KIN>

F:10-18/Region: protein kinase ATP-binding motif

F:14,16/Binding site: phosphate (Thr) (covalent) #status predicted

F:15/Binding site: phosphate (Tyr) (covalent) #status predicted

F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 10.6%, Score 233, DB 2, Length 298;

Best Local Similarity 29.0%, Pred. No. 1,46-06;

Matches 98, Conservative 43, Mismatches 119; Indels 78; Gaps 16;

QY 23 WQVSRSGSGSSAVYRVRCGNGSPFGALKEF-LPGGTGAASAAYGFRKERALE 81

Db 4 FQKYEKIGEGYGVYKAR--NKLTEGVYALKRLDLETBGSVSTAI-----REISLKK 56

QY 82 QLGCHRNIVTLGYVFTIHSPNVSRCLLLELDVSVSELLYSHQGSGMMIQCARD 141

Db 57 EL-NHPNIVKLLDV--IHTENKL--YLVEFFLLHQDKEKFDADALGIPLEPLIKSYLFQ 110

QY 142 VLEMLAFHNGGVYHADLKPNNILMSAENECFKLIDEGLSPEKGNQVKK---IQTDGVR 198

Db 111 LLQGLAFCHSHRVLARDLKPQNLILNTEG-AIKLADDEGLAANAFVPRVTTIHVEYVTLMYR 169

QY 199 APEAELONCIAQAGLQSDTECTSAVDLMSLGIILLEMFSGMKLGHTVRSQEWKANSAAI 258

Db 170 APEI-LIGC-----KYSTAVDIMSIGCTFAEM-----VTRRALFPEDSE--I 209

QY 259 DHFASAKAVN-----AATPAYHL-----RDLIKSMLHD 287

Db 210 DQLRFRIRLTGTPDEVVWPGVTSMPDYKPSFPKARODFSKVVPDLDEGRSLLSQMLHY 269

QY 288 DPSRRIPAEMLGSPFSPFPAPRIEDLVMLPTVLR 325

Db 270 DPNKRISKALAHPIF-----QD-VTPQVPHKL 298

RESULT 13

A56492

protein kinase ERK2 (EC 2.7.1.-) - slime mold (Dictyostelium discoideum)

C:Species: Dictyostelium discoideum

C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004

C:Accession: A56492

C:Segall, J.E.; Knappe, A.; Shaulsky, G.; Ecker, M.; Maeda, M.; Gaskins, C.; Firtel, R.A.

J. Cell Biol. 128, 405-413, 1995
A>Title: A MAP kinase necessary for receptor-mediated activation of adenylyl cyclase in
A.Reference number: A56492; MUID:95146547; PMID:7644154
A.Accession: A56492
A>Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-369 <SEG>
C/Cross-references: UNIPROT:Q7M445; GB:L33043
C/Superfamily: kinase-related transforming protein; protein kinase homology
C/Keywords: ATP; phosphotransferase
F/12-273/Domain: protein kinase homology <KIN>
F/20-28/Region: protein kinase ATP-binding motif

Query Match 10.6%; Score 233; DB 2; Length 369;
Best Local Similarity 26.9%; Pred. No. 1,7e-06;
Matches 91; Conservative 52; Mismatches 109; Indels 86; Gaps 17;

Oy 66 ASAAEYFKERPALELOGHNRIVTLVGYFTTHFSPPNVRCLLLELDVSSEL-----121
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 52 ATQAQRFR-EIMFELGHEHNIKLNVTK--ADNDRIYLVEHHETDLAAVIRAK 107
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Oy 122 LTVSSHOGCGMMWIOHCARDVLEALFLHHEGVHADLKPRNLMSAENECEF-KTIDFGI 180
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 108 ILSEIKR-----QITTYGLKALKKMHSANVLHRDIKSNTL--LNSCLVKVADFGL 158
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Oy 181 -----SFEKNQDV-KYIQDGYRAPEALONCLAQGLOSTECTSAVDLSLGIT 231
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 159 ARSITSLESIAENNPVTEVATRWYAPRI-----LIGSKYTKGVDMWSIGCI 208
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Oy 232 LLE-----MEFG-----MKLKHTRSGEMWANSAILIHIFASAVVN-----269
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 209 LGHELGEKANPFGNSTMNQLILIEVTGRPSADIEAIKSPFGTMLESIPSPNPSISLD 268
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Oy 270 ----AAIPAYHLRLDKSMHDPPSRRIPEAMALCSPF---FSIPF-APHIEDLVMLPT- 320
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 269 MYPSASVDAL---DLKKGSQFNPDKRITAEALAHAFVLYGFHNPAEBPHDRILIKSID 325
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Oy 321 ----PYRLNLNVDDDYLGNEEEYEDVEDVDEECOK 353
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 326 DGQKFPIAEYRNRL-----YNDIIKKKEEKREK 353
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :

RESULT 14
T133024
probable protein kinase (EC 2.7.1.-) F8L21.120 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jan-2000
C/Accession: T133024
R/Beyan, M.; Peters, S.A.; van Steveren, M.; Dirke, W.; Stiekema, W.; Bancroft, I.; New
submitted to the Protein Sequence Database, July 1999
A/Reference number: Z17587
A/Accession: T133024
A/Molecule type: DNA
A/Residues: 1-373 <BEV>
A/Cross-references: EMBL:AL096882; GSPDB:GN00062; ATSP:F8L21.120
A/Experimental source: cultivar Columbia; BAC clone F8L21
C/Genetics:
A/Gene: ATSP:F8L21.120
A/Map position: 4
A/Intons: 58/1; 98/3; 144/3; 255/3; 317/1
C/Superfamily: kinase-related transforming protein; protein kinase homology
C/Keywords: phosphotransferase; protein kinase
F/40-295/Domain: protein kinase homology <KIN>

Query Match 10.4%; Score 230; DB 2; Length 373;
Best Local Similarity 25.9%; Pred. No. 2,6e-06;
Matches 83; Conservative 59; Mismatches 111; Indels 68; Gaps 14;

Oy 75 KERAALBOLGHRNIVTLVGYFTTHFSPPNVRCLLLELDVSSELITYSSHQCGSMW 134
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 86 REIKLRHLE-HENVAVIKIIRPKKEDPVYVIYFELMDTDLHQIT--RSNGSLNDH 142
 : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Oy 135 IOHCARDVLEALFLHHEGVHADLKPRNLMSAENECEFKTIDFGL-SFEKNQDVKYI 192

| | | | | |
|----|-----|---|---|-----|
| Db | 143 | QYFLQYLTIRGKXHYHSANVLRHDLKPSNLLNS | NCDELKIDPGLARTTSSETEMYTEYV | 201 |
| QY | 193 | QTDGTRAPAEALQNCIAQGLQSDTECTSAVDLMSGIIILE | -----MFGSKLKYTR | 246 |
| Db | 202 | VTWRVRAPEL | -----LINSSTYTAIDVMSVGCIPAEIMTREPPEPGDYHQK | 251 |
| QY | 247 | -----SOEW--KANSSAIDHI | -----FASK--AVVNAAIPTVYHLRLDILS | 283 |
| Db | 252 | LITTELGSPDGASLEFLRSGANARKYKEIPKPRQNFSAFSPMSNSTAI | -----DLLEK | 305 |
| QY | 284 | MLHDPSRRIRPAMMLCSPFSPISIPAPHIEDLVMLEPTVRLRLNLTAVDDVDLGNBEEDEV | | 343 |
| Db | 306 | MLVFPPVKRITVEALNLCYPYLSA | -----LHLDNDEP | 352 |
| QY | 344 | VEDVEE-----ECQKIGPVVSL | 360 | |
| Db | 353 | EEIKELVWLSEVKRNFPLPSI | 373 | |

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 122.5 Seconds
(without alignments)

1968.016 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206

Sequence: 1 MAGSGCAGCAEPPRFLEAFG.....VATFYPLSAVKGKYLQTL 419

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 2198 | 99.6 | 419 | 1 | KIST_HUMAN |
| 2 | 2186 | 99.1 | 419 | 1 | KIST_RAT |
| 3 | 2185 | 99.0 | 419 | 2 | AAH58732 |
| 4 | 2185 | 99.0 | 419 | 2 | AAO13515 |
| 5 | 2180 | 98.8 | 419 | 1 | KIST_MOUSE |
| 6 | 1711 | 77.6 | 330 | 2 | O8C0N6 |
| 7 | 257 | 11.7 | 505 | 2 | O6CFX3 |
| 8 | 248.5 | 11.3 | 882 | 2 | O8M0L7 |
| 9 | 243.5 | 11.0 | 395 | 2 | O6GM90 |
| 10 | 243.5 | 11.0 | 720 | 2 | O9CA22 |
| 11 | 243.5 | 11.0 | 765 | 2 | O25758 |
| 12 | 243.5 | 11.0 | 777 | 2 | O84VX4 |
| 13 | 243.5 | 11.0 | 826 | 2 | O94656 |
| 14 | 243.5 | 11.0 | 914 | 2 | O81IF0 |
| 15 | 242.5 | 11.0 | 395 | 2 | O84YR3 |
| 16 | 242.5 | 11.0 | 395 | 2 | AAH70745 |
| 17 | 242 | 11.0 | 433 | 2 | O84SN3 |
| 18 | 241.5 | 10.9 | 387 | 1 | SKPI_SCHPO |
| 19 | 240.5 | 10.9 | 360 | 2 | O6S267 |
| 20 | 240.5 | 10.9 | 360 | 2 | AAH17088 |
| 21 | 240 | 10.9 | 368 | 1 | SMKI_YEAST |
| 22 | 238 | 10.8 | 288 | 1 | CDK2_MESAU |
| 23 | 235.5 | 10.7 | 367 | 1 | MAK1_RAT |
| 24 | 235.5 | 10.7 | 434 | 2 | O75B85 |
| 25 | 235.5 | 10.7 | 434 | 2 | AA551605 |
| 26 | 234.5 | 10.6 | 367 | 1 | MAK1_MOUSE |
| 27 | 234.5 | 10.6 | 367 | 2 | O6ZD93 |
| 28 | 234.5 | 10.6 | 367 | 2 | BAC99508 |
| 29 | 234.5 | 10.6 | 433 | 1 | KKQ1_YEAST |
| 30 | 234 | 10.6 | 298 | 1 | CDK2_CRIGR |
| 31 | 234 | 10.6 | 298 | 1 | CDK2_RAT |

| | | | | | | |
|----|-------|------|------|---|------------|---------------------|
| 32 | 234 | 10.6 | 298 | 2 | O6P751 | O6P751 rattus norv |
| 33 | 234 | 10.6 | 298 | 2 | AAH61832 | AAH61832 rattus no |
| 34 | 234 | 10.6 | 361 | 1 | JNK SUBDO | O966Y3 subteriles d |
| 35 | 234 | 10.6 | 435 | 1 | HOG1_YEAST | P32485 saccharomyc |
| 36 | 233.5 | 10.6 | 433 | 2 | O6B629 | O6B629 parametium |
| 37 | 233.5 | 10.6 | 1233 | 2 | O6CAC2 | O6CAC2 yarrowia li |
| 38 | 233 | 10.6 | 298 | 1 | CDK2_CARAU | P43450 carassius a |
| 39 | 233 | 10.6 | 298 | 1 | CDK2_HUMAN | P24941 homo sapien |
| 40 | 233 | 10.6 | 298 | 2 | AAH35467 | AAH35467 homo sapi |
| 41 | 233 | 10.6 | 369 | 2 | O7M445 | O7M445 dictyosteli |
| 42 | 233 | 10.6 | 392 | 2 | O9YMAN3 | O9YMAN3 drosophila |
| 43 | 232.5 | 10.5 | 357 | 2 | O9HD31 | O9HD31 homo sapien |
| 44 | 232.5 | 10.5 | 385 | 2 | O8IU85 | O8IU85 homo sapien |
| 45 | 232.5 | 10.5 | 385 | 2 | O8BW96 | O8BW96 mus musculu |

ALIGNMENTS

RESULT 1
KIST_HUMAN STANDARD; PRT; 419 AA.
ID KIST_HUMAN Q8TAS1; Q96C22;
AC Q8TAS1; Q96C22;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting with statminin).
GN Name=KIST; Synonyms=KIS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX PubMed=12782393;
RA Bieche I., Manoeuvre V., Curmi P.A., Laurendeau I., Lachkar S., Leroy K., Vidaud D., Sobel A., Maucuer A.;
RT "Quantitative RT-PCR reveals a ubiquitous but preferentially neural expression of the KIS gene in rat and human."
RL Brain Res. Mol. Brain Res. 114:55-64(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Eye, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo W.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman W., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Possesses kinase activity. May be involved in trafficking and/or processing of RNA (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with statmin (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;

```

CC      IcoId=Q8TAS1-1; Sequence=Displayed;
CC      Name=2;
CC      IcoId=Q8TAS1-2; Sequence=VSP_004908, VSP_004909;
CC      Note=No experimental confirmation available;
CC      SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC      -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AT536197; CAD60192.1; -.
DR      EMBL; BC014917; AAH14917.1; -.
DR      EMBL; BC026046; AAH26046.1; -.
DR      HSSP; P26368; 100P.
DR      GO; GO:0005737; Cytoplasm; ISS.
DR      GO; GO:0005634; C:nucleus; ISS.
DR      GO; GO:0005524; F:ATP binding; ISS.
DR      GO; GO:0005515; F:protein binding; ISS.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR      GO; GO:0003723; F:RNA binding; ISS.
DR      GO; GO:0018105; P:pepckidyl-serine phosphorylation; ISS.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR000719; Prot kinase.
DR      InterPro; IPR000504; RNA Rec mot.
DR      InterPro; IPR008271; Ser_Thr_kin_AS.
DR      Pfam; PF00669; Kinase; 1.
DR      Pfam; PF00076; RRM_1; 1.
DR      ProDom; PD000001; Prot kinase; 1.
DR      SMART; SM00360; RRM; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR      PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00109; PROTEIN_KINASE_ST; FALSE_NEG.
DR      PROSITE; PS00102; RRM; 1.
DR      PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR      KMW Alternative splicing; ATP-binding; Nuclear protein; RNA-binding;
DR      Serine/threonine-protein kinase; Transferase.
FT      DOMAIN 23 304 Protein kinase.
FT      DOMAIN 324 406 RNA-binding (RRM).
FT      NP_BIND 29 37 ATP (By similarity).
FT      BINDING 54 54 ATP (By similarity).
FT      ACT_SITE 158 158 Proton acceptor (By similarity).
FT      VARSPLIC 342 344 DVV -> GIC (in isoform 2).
FT      VARSPLIC 345 419 Missing (in isoform 2).
SQ      SEQUENCE 419 AA; 46546 MW; 903E982BE12A8C68 CRC64;
Query Match 99.6%; Score 2198; DB 1; Length 419;
Best Local Similarity 99.8%; Pred. No. 1.3e-159;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

DB      241 LKHTVRSQEWKANSALIDHIFASKAVVNAAPAYHLRLIKSLHDDPSRRIPEMALC 300
QY      301 SPFFSIPFAPHIEDLMLPTPYRLNLVADDYLGNEEYEDVEDVEKCECKYGPVSL 360
DB      301 SPFFSIPFAPHIEDLMLPTPYRLNLVADDYLGNEEYEDVEDVEKCECKYGPVSL 360
QY      361 LVPKENPGQGVFEVEYANAGDSKAAQKLLTGMPDGKPVATFPYPLSAVRYQTLL 419
DB      361 LVPKENPGQGVFEVEYANAGDSKAAQKLLTGMPDGKPVATFPYPLSAVRYQTLL 419
RESULT 2
ID      KIST_RAT STANDARD; PRT; 419 AA.
AC      063285;
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE      with statmin) (PAM COOH-terminal interactor protein 2) (P-CIP2).
GN      Name=Kist; Synonyms=Kfs;
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX      NCBI_TaxId=10116;
RN      [1]
RP      SEQUENCE FROM N.A., FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, TISSUE
RP      SPECIFICITY, AND MUTAGENESIS OF LYS-54.
RC      TISSUE=Striatum;
RX      MEDLINE=97435279; PubMed=9287318;
RA      Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
RA      Sobel A.;
RT      "Kis is a protein kinase with an RNA recognition motif.";
RL      J. Biol. Chem. 272:23151-23156(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Caldwell B.D., Darlington D.N., Penzes P., Johnson R.C., Bipper B.A.,
RA      Maine R.E.;
RL      Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 28-419 FROM N.A.
RC      TISSUE=Hippocampus;
RX      MEDLINE=97067094; PubMed=8910496;
RA      Alam R., Caldwell B.D., Johnson R.C., Darlington D.N., Maine R.E.,
RA      Bipper B.A.;
RT      "Novel proteins that interact with the COOH-terminal cytosolic routing
RT      determinants of an integral membrane peptide-processing enzyme.";
RL      J. Biol. Chem. 271:28636-28640(1996).
CC      -1- FUNCTION: Possesses kinase activity. May be involved in
CC      trafficking and/or processing of RNA.
CC      -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC      -1- SUBUNIT: Interacts with statmin (By similarity).
CC      -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC      -1- TISSUE SPECIFICITY: In the embryo, preferentially expressed in the
CC      developing nervous system.
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC      -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X98374; CAA67021.1; -.
DR      EMBL; U70372; AAC53031.2; -.
DR      HSSP; P26368; 100P.
DR      RGD; 2968; Kistf.
DR      GO; GO:0005737; Cytoplasm; IDA.
DR      GO; GO:0005634; C:nucleus; IDA.

```

DR GO; GO:000524; F.ATP binding; IC.
DR GO; GO:000515; F.protein binding; ISS.
DR GO; GO:0004674; F.protein serine/threonine kinase activity; IDA.
DR GO; GO:0003723; F.RNA binding; NAS.
DR GO; GO:0018105; P.peptidyl-serine phosphorylation; IDA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR005054; RNA rec. mod.
DR InterPro; IPR008271; Ser-thr_pkin_AS.
DR Pfam; PF00069; PKinase; 1.
DR Pfam; PF00076; RRM_1; 1.
DR ProDom; PDO00001; Prot_kinase; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS00102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KM ATP-binding; Nuclear protein; RNA-binding;
Serine/threonine-protein kinase; Transferase.
FT DOMAIN 23 304 Protein kinase.
FT NP_BIND 324 406 RNA-binding (RRM).
FT BINDING 29 37 ATP (By similarity).
FT ACT_SITE 141 141 Proton acceptor (By similarity).
FT ACT_SITE 158 158 Proton acceptor (By similarity).
FT MUTAGEN 54 54 K->R: Loss of activity.
SQ SEQUENCE 419 AA; 46505 MW; D2F9B7BF8F080DE3 CRC64;

Query Match 99.1%; Score 2186; DB 1; Length 419;
Best Local Similarity 99.0%; Pred. No. 1.1e-158;
Matches 415; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPRPLEAFGRMLQVOSRLGSSSSASVYRVRCGPGSPGALKOFLPPG 60
DB 1 MAGSGCAGWAEPRPLEAFGRMLQVOSRLGSSSSASVYRVRCGPGSPGALKOFLPPG 60
QY 61 TTGAASAASAEYGFRRERAALEOLQGHRIVTLYGVFTIHSPVPSRCILLETLDVSVSE 120
DB 61 TTGAASAASAEYGFRRERAALEOLQGHRIVTLYGVFTIHSPVPSRCILLETLDVSVSE 120
QY 121 LLLYSSHOGCSMMMIQHCARDVLEALFLHHEGYVHADLKPRNIIIMSAENECFKLIDFGL 180
DB 121 LLLYSSHOGCSMMMIQHCARDVLEALFLHHEGYVHADLKPRNIIIMSAENECFKLIDFGL 180
QY 121 LLLYSSHOGCSMMMIQHCARDVLEALFLHHEGYVHADLKPRNIIIMSAENECFKLIDFGL 180
DB 121 LLLYSSHOGCSMMMIQHCARDVLEALFLHHEGYVHADLKPRNIIIMSAENECFKLIDFGL 180
QY 181 SFKEGNQDVYKIQTGYPAPAEIQLNCLAQAGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
DB 181 SFKEGNQDVYKIQTGYPAPAEIQLNCLAQAGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
QY 241 LKHTVRSQEWKANSALIDHIFASKAVVNAIIPAYHLRDLIKSLMDPERRIPAEMLC 300
DB 241 LKHTVRSQEWKANSALIDHIFASKAVVNAIIPAYHLRDLIKSLMDPERRIPAEMLC 300
QY 241 LKHTVRSQEWKANSALIDHIFASKAVVNAIIPAYHLRDLIKSLMDPERRIPAEMLC 300
DB 241 LKHTVRSQEWKANSALIDHIFASKAVVNAIIPAYHLRDLIKSLMDPERRIPAEMLC 300
QY 301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVEDVKECCQKYGPVVSL 360
DB 301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVEDVKECCQKYGPVVSL 360
QY 361 LVPKENPGRGQVFEVYANAGDSKAAQKLLTGEMFGKPVVATFPLSAVKGRLYQTL 419
DB 361 LVPKENPGRGQVFEVYANAGDSKAAQKLLTGEMFGKPVVATFPLSAVKGRLYQTL 419

RESULT 3
AAH58732
ID AAH58732 PRELIMINARY; PRT; 419 AA.
AC AAH58732;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Kinase interacting with leukemia-associated gene (Statmin).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22386257; PubMed=12477932;
RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uebli T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Mair M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RI Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strusberg R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC058732; AAH58732.1; -.
KW Kinase.

Query Match 99.0%; Score 2185; DB 2; Length 419;
Best Local Similarity 99.0%; Pred. No. 1.3e-158;
Matches 415; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPRPLEAFGRMLQVOSRLGSSSSASVYRVRCGPGSPGALKOFLPPG 60
DB 1 MAGSGCAGWAEPRPLEAFGRMLQVOSRLGSSSSASVYRVRCGPGSPGALKOFLPPG 60
QY 61 TTGAASAASAEYGFRRERAALEOLQGHRIVTLYGVFTIHSPVPSRCILLETLDVSVSE 120
DB 61 TTGAASAASAEYGFRRERAALEOLQGHRIVTLYGVFTIHSPVPSRCILLETLDVSVSE 120
QY 61 TTGAASAASAEYGFRRERAALEOLQGHRIVTLYGVFTIHSPVPSRCILLETLDVSVSE 120
DB 61 TTGAASAASAEYGFRRERAALEOLQGHRIVTLYGVFTIHSPVPSRCILLETLDVSVSE 120
QY 121 LLLYSSHOGCSMMMIQHCARDVLEALFLHHEGYVHADLKPRNIIIMSAENECFKLIDFGL 180
DB 121 LLLYSSHOGCSMMMIQHCARDVLEALFLHHEGYVHADLKPRNIIIMSAENECFKLIDFGL 180
QY 121 LLLYSSHOGCSMMMIQHCARDVLEALFLHHEGYVHADLKPRNIIIMSAENECFKLIDFGL 180
DB 121 LLLYSSHOGCSMMMIQHCARDVLEALFLHHEGYVHADLKPRNIIIMSAENECFKLIDFGL 180
QY 181 SFKEGNQDVYKIQTGYPAPAEIQLNCLAQAGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
DB 181 SFKEGNQDVYKIQTGYPAPAEIQLNCLAQAGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
QY 241 LKHTVRSQEWKANSALIDHIFASKAVVNAIIPAYHLRDLIKSLMDPERRIPAEMLC 300
DB 241 LKHTVRSQEWKANSALIDHIFASKAVVNAIIPAYHLRDLIKSLMDPERRIPAEMLC 300
QY 241 LKHTVRSQEWKANSALIDHIFASKAVVNAIIPAYHLRDLIKSLMDPERRIPAEMLC 300
DB 241 LKHTVRSQEWKANSALIDHIFASKAVVNAIIPAYHLRDLIKSLMDPERRIPAEMLC 300
QY 301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVEDVKECCQKYGPVVSL 360
DB 301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVEDVKECCQKYGPVVSL 360
QY 361 LVPKENPGRGQVFEVYANAGDSKAAQKLLTGEMFGKPVVATFPLSAVKGRLYQTL 419
DB 361 LVPKENPGRGQVFEVYANAGDSKAAQKLLTGEMFGKPVVATFPLSAVKGRLYQTL 419

RESULT 4
AAO13515
ID AAO13515 PRELIMINARY; PRT; 419 AA.
AC AAO13515;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)

DT 02-MAR-2004 (TEMBLrel. 27, Last annotation update)
 DE KIS kinase.
 GN KIS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBL_taxid=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=12956;
 RL Crook M.F., Boehm M., Nabel E.G.;
 DR Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
 KW NCBL: AY180177; AA013515.1; -
 SQ SEQUENCE 419 AA; 46489 MW; 3BD3C06A59E22540 CRC64;
 Query Match 99.0%; Score 2185; DB 2; Length 419;
 Best Local Similarity 99.0%; Pred. No. 1.3e-158;
 Matches 415; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MAGSGCAGMAGPPRFLFAFGRLLWQVOSRLGSGSASVYRVRCGNGSPGALKQFLPPG 60
 DB 1 MAGSGCAGMAGPPRFLFAFGRLLWQVOSRLGSGSASVYRVRCGNGSPGALKQFLPPG 60
 QY 61 TTGAAASAAATYGRKRAALAEQLQGHNITLYGVFTIHSPVPSKCLLELDVSVSE 120
 DB 61 TTGAAASAAATYGRKRAALAEQLQGHNITLYGVFTIHSPVPSKCLLELDVSVSE 120
 QY 121 LLLYSSHQGSMWMIQHCARDVLEALFLHHEGVHADLPRIILMSAENECFLIDFGL 180
 DB 121 LLLYSSHQGSMWMIQHCARDVLEALFLHHEGVHADLPRIILMSAENECFLIDFGL 180
 QY 181 SFPEKQDVYKIQTDGYRAPEALONCLAQGLQSDTECTSAVDLWSLGIILLEMESGMK 240
 DB 181 SFPEKQDVYKIQTDGYRAPEALONCLAQGLQSDTECTSAVDLWSLGIILLEMESGMK 240
 QY 241 LKRTVNSQEKANSSAIIIDHIFASKAVVNAAPYAHRLDKSMHDDPSRRIPAEALC 300
 DB 241 LKRTVNSQEKANSSAIIIDHIFASKAVVNAAPYAHRLDKSMHDDPSRRIPAEALC 300
 QY 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVLDLDDYLGNEEEDVEDVEKESCKYGPVSL 360
 DB 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVLDLDDYLGNEEEDVEDVEKESCKYGPVSL 360
 QY 361 LVPEKENGRCQVFEVYANAGDSKAAQGLTGRMFQKFFVATFPYPSAYKGYLYQTL 419
 DB 361 LVPEKENGRCQVFEVYANAGDSKAAQGLTGRMFQKFFVATFPYPSAYKGYLYQTL 419
 RESULT 5
 KIST MOUSE STANDARD; PRT; 419 AA.
 ID KIST MOUSE STANDARD; PRT; 419 AA.
 AC P97343; Q61775; Q9CYTL;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting with statmin).
 GN Name=Kist; Synonym=KIS;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBL_taxid=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=22354683; PubMed=1246851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kaaukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Oosato N., Saito R., Suzuki H., Yamataka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Baldirelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kampin A., Matsuda H., Batalov S., Batzel K.W.,
 RA Blake U.A., Bradt D., Bruste V., Chochia C., Cordani L.B., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guenrich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kanagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglocz D.R., Malraiz L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perera G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempé C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wallestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyszewski B., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 RN [2]
 RP SEQUENCE OF 1-414 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97435279; PubMed=9287318;
 RA Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
 RA Sobel A.;
 RT "KIS is a protein kinase with an RNA recognition motif."
 RL J. Biol. Chem. 272:23151-23156(1997).
 RN [3]
 RP REVISIONS TO 52 AND 68-69.
 RA Maucuer A.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 122-321 FROM N.A., AND INTERACTION WITH STATMIN.
 RC TISSUE=Embryo;
 RX MEDLINE=95241452; PubMed=7724523;
 RA Maucuer A., Camonis J.H., Sobel A.;
 RT "Statmin interaction with a putative kinase and cold-coil-forming
 RT protein domains."
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3100-3104(1995).
 CC -1- FUNCTION: Possesses kinase activity. May be involved in
 CC trafficking and/or processing of RNA (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated
 CC -1- SUBUNIT: Interacts with statmin.
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Y10725; CAA71714.2; -
 DR EMBL: AK013347; BAB28802.1; -
 DR EMBL: X62320; CAA57763.1; -
 DR PIR: I48615; I48615.
 DR HSSP: P26368; 100P.
 DR MGD: MGI:1341908; Kist.
 DR GO: GO:0005737; Cytoplasm; ISS.
 DR GO: GO:0005634; Cytoplasm; ISS.
 DR GO: GO:0005524; F-ATP binding; ISS.
 DR GO: GO:0005515; F-protein binding; IPT.
 DR GO: GO:0004674; F-protein serine/threonine kinase activity; ISS.
 DR GO: GO:0003723; F-RNA binding; ISS.
 DR GO: GO:0018105; Peptidyl-tyrosine phosphorylation; ISS.
 DR InterPro: IPR011009; Kinase_like.

DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR000504; RNA rec mot.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR Pfam; PF00076; RRM_1; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 DR ATP-binding; Nuclear Protein; RNA-binding;
 KM Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 23 303 Protein kinase.
 FT NE_BIND 323 405 RNA-binding (RRM).
 FT BINDING 29 37 ATP (By similarity).
 FT ACT_SITE 54 54 ATP (By similarity).
 FT CONFLICT 158 158 Proton acceptor (By similarity).
 FT CONFLICT 170 170 N -> D (in Ref. 4).
 FT CONFLICT 231 231 T -> I (in Ref. 3 and 4).
 SQ SEQUENCE 419 AA; 46477 MW; A91B307BA50F29A1 CRC64;

Query Match 98.8%; Score 2180; DB 1; Length 419;
 Best Local Similarity 98.8%; Pred. No. 3.2e-158;
 Matches 414; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MASSGCMWAEPPRFLEAFGRFLNQVSRSGSSASTYRRRCGNGPSPGALKQFLPPG 60
 1 MASSGCMWAEPPRFLEAFGRFLNQVSRSGSSASTYRRRCGNGPSPGALKQFLPPG 60
 DB 1 TTTGAASAAEYGRKERAALEQOGHNITLYGVFTIHSPVNSCLLELLDVSE 120
 61 TTTGAASAAEYGRKERAALEQOGHNITLYGVFTIHSPVNSCLLELLDVSE 120
 QY 121 LLLYSSHOQSMWMIQHCARDVLEALFLHREGVHADLPRNIIILSAENECEFLDPGL 180
 121 LLLYSSHOQSMWMIQHCARDVLEALFLHREGVHADLPRNIIILSAENECEFLDPGL 180
 DB 122 LLLYSSHOQSMWMIQHCARDVLEALFLHREGVHADLPRNIIILSAENECEFLDPGL 180
 122 LLLYSSHOQSMWMIQHCARDVLEALFLHREGVHADLPRNIIILSAENECEFLDPGL 180
 QY 181 SPFEGNDVYKTIQDTGRAPAEALQNLQAGLOSTRECTSANDVMSLGIILLEMPSGMK 240
 181 SPFEGNDVYKTIQDTGRAPAEALQNLQAGLOSTRECTSANDVMSLGIILLEMPSGMK 240
 DB 181 SPFEGNDVYKTIQDTGRAPAEALQNLQAGLOSTRECTSANDVMSLGIILLEMPSGMK 240
 181 SPFEGNDVYKTIQDTGRAPAEALQNLQAGLOSTRECTSANDVMSLGIILLEMPSGMK 240
 QY 241 LKHTVRSQEWKANSATIIHIFASKAVYNAIIPYHLRLIKSLHDDPSRRIPAEVALC 300
 241 LKHTVRSQEWKANSATIIHIFASKAVYNAIIPYHLRLIKSLHDDPSRRIPAEVALC 300
 DB 241 LKHTVRSQEWKANSATIIHIFASKAVYNAIIPYHLRLIKSLHDDPSRRIPAEVALC 300
 241 LKHTVRSQEWKANSATIIHIFASKAVYNAIIPYHLRLIKSLHDDPSRRIPAEVALC 300
 QY 301 SPFFSIPFAHIEDLVMLPTPVRLINLVDDYLGNEEYEDVEDYKECCQKGPVSL 360
 301 SPFFSIPFAHIEDLVMLPTPVRLINLVDDYLGNEEYEDVEDYKECCQKGPVSL 360
 DB 301 SPFFSIPFAHIEDLVMLPTPVRLINLVDDYLGNEEYEDVEDYKECCQKGPVSL 360
 301 SPFFSIPFAHIEDLVMLPTPVRLINLVDDYLGNEEYEDVEDYKECCQKGPVSL 360
 QY 361 LVPEKENGQGVFEVYANAGDSKAAQKLTIGRMFDGKFVATVPYPLSAVRYGYQTLL 419
 361 LVPEKENGQGVFEVYANAGDSKAAQKLTIGRMFDGKFVATVPYPLSAVRYGYQTLL 419
 DB 361 LVPEKENGQGVFEVYANAGDSKAAQKLTIGRMFDGKFVATVPYPLSAVRYGYQTLL 419
 361 LVPEKENGQGVFEVYANAGDSKAAQKLTIGRMFDGKFVATVPYPLSAVRYGYQTLL 419

RESULT 6
 Q8CON6 PRELIMINARY; PRT; 330 AA.
 ID Q8CON6
 AC Q8CON6;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
 DE library, clone:493401D07 product:kinase interacting with leukemia-
 DE associated gene (statmin), full insert sequence.
 GN Name=Kist;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Mech. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN PANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA The FANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=2049374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Ozawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ogawa K., Tanaka T., Marsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takada Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBD databases.
 DR EMBL; AK030152; BAC26812.1; -.
 DR HSSP; P26368; 100P.
 DR MGD; MGI:1341908; Kist.
 DR GO; GO:000524; F-ATP binding; IEA.
 DR GO; GO:0004672; F-protein kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00069; Pkinase; 2.
 DR Pfam; PF00076; RRM_1; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

| DR | PROSITE; PS50102; RRM; 1. |
|----|--|
| KW | LINEAGE |
| SC | SEQUENCE 330 AA; 37201 MW; 4164C0758BE7DB4C CRC64; |
| QY | Query Match 77.6%; Score 1711; DB 2; Length 330; Best Local Similarity 98.8%; Pred. No. 1.7e-122; |
| Db | Matches 326; Conservative 2; Mismatches 2; Indels 0; Gaps 0 |
| QY | 90 VTLGVFTIHSPNVPSCRLLLELDVSVSELLLYSHQCGSMWIOHCARDVLEALFL 149 : |
| Db | 1 MTLGVFTIHSPNVPSCRLLLELDVSVSELLLYSHQCGSMWIOHCARDVLEALFL 60 |
| QY | 150 HHEGVVHADLPRLNLSAENECFLIDFGSFKSGNDVKKIORDGRAPAEI.ONCIA 209 |
| Db | 61 HHEGVVHADLPRLNLSAENECFLIDFGSFKSGNDVKKIORDGRAPAEI.ONCIA 120 |
| QY | 210 QAGLSDPECTSAVDLMSLGIIILTEMFGSMKLTIVRSQEWKANSATIIDHIFASKAVVN 269 |
| Db | 121 QAGLSDPECTSAVDLMSLGIIILTEMFGSMKLTIVRSQEWKANSATIIDHIFASKAVVN 180 |
| QY | 270 AAIPTAYHLRDLIKSMLHDDPSRRIPAEMLCSPPFSIPPAHIEDLVMLPTVLRLLNVL 329 |
| Db | 181 AAIPTAYHLRDLIKSMLHDDPSRRIPAEMLCSPPFSIPPAHIEDLVMLPTVLRLLNVL 240 |
| QY | 330 DDDVLGNSEEEEDVEDVEDEKCEQKGPVSLVLPKENGRCGVPFVEYVANAAGSKAAOKLL 389 |
| Db | 241 DDDVLGNSEEEEDVEDVEDEKCEQKGPVSLVLPKENGRCGVPFVEYVANAAGSKAAOKLL 300 |
| QY | 390 TGRMFDEGKVVATFYPPLSAYKRGVLYQTLL 419 |
| Db | 301 TGRMFDEGKVVATFYPPLSAYKRGVLYQTLL 330 |

RESULT 7

Q6CFX3 PRELIMINARY; PRT: 505 AA.

Q6CFX3
AC Q6CFX3;
DT 01-OCT-2004 (TREMBLrel. 28, Created)
DT 01-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Chromosome B of strain CL1899 of Yarrowia lipolytica.
GN ORFNames=YALI0B02816g;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OC NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG GENOLIVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Mack C., Neuveglise C., Talla E.,
RA Goffard N., Prangeul L., Aigle M., Anthonard V., Babour A., Barbe V.,
RA Barney S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boistrame A., Boyer J., Catcolico U., Confanetra F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dunazet H., Groppi A.,
RA Hantreya F., Hemequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerret A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Micaud J.M., Nikolski M., Ozias S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potter S., Richard G.F., Strub M.L., Suleau A.,
RA Swenne D., Tekla F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeilou-Meyer M., Zivanovic I., Boletín-Pukhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Galliard C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts."
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genoscope;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL, CR382128; CAG82657.1; --
Q6 SEQUENCE 505 AA; 58191 MW; D30933013FCE357 CRC64;

| | | |
|----|---|---|
| | Query Match | 11.7%; Score 257; DB 2; Length 505; |
| | Best Local Similarity | 25.5%; Pred. No. 4.5e-11; |
| | Matches | 88; Conservative 46; Mismatches 129; Indels 82; Gaps 11 |
| Qy | 23 WOVQSRLGSGSSAVYRVRCGGNGSPGALKOFLPPTGTGAASAAEYGRKREALLAQ | 82 |
| Dd | 22 FQVTKELGHAGVGIVCAAKYTGTDTDGVATIKVTNIPSKNILCKRA---LRELKLNH | 77 |
| Qy | 83 LQGHRIIVLYGVTTIHPSNVPRCLLELLDV-SVSELLLSSHGCCSMWMT----- | 135 |
| Dd | 78 FRGHKNITCLVD-----MDIVDPNNFNEELYELBELMCDMHQIIRSQP | 121 |
| Qy | 136 -----QHCAADVLEALFLHHEGYVHADLPKRNILMSAEKCFLLIDFGSFKEGNOY | 189 |
| Dd | 122 LTDAHYOSFYVQLLAGVKTIHSADVLHRDLRPGMLVNVADE-LKICDFGLARGSDD | 180 |
| Qy | 190 K-----YICTDGYRAPAEALQNCLAQAQIGSDTECTSAVDLSLGILLMEFSGMKLK | 242 |
| Dd | 181 KNAGFLTEYVATRMYRAPETWL-----STQSYTKAIDIMSVGCILABELGG---K | 227 |
| Qy | 243 HTVRSGEWKANSSAIIDHI-----FAASKVVNNAALPAYH-- | 276 |
| Dd | 228 PLFGKKNYVVOALNIHLVLTGPSEETLKRGSPPAOEYVRGLPPMPKP.PFSTLPEFPANPE | 287 |
| Qy | 277 LRDLIKSLMHDDPRSRI.PAEMALCSPPESIPFAFHIEDLVNLPPP | 321 |
| Dd | 288 ALDLIELKMLAFDPERAYVEBALHPHYUKTWHD--RDREPACFTP | 330 |

| RESULT # | Q8MQL7 | PRELIMINARY | PRT | 882 AA. |
|----------|--|-------------|-----|---------|
| ID | Q8MQL7 | | | |
| AC | Q8MQL7; Q22155; | | | |
| DT | 01-MAR-2002 (TrEMBLrel. 20, Created) | | | |
| DT | 01-MAR-2002 (TrEMBLrel. 20, Last sequence update) | | | |
| DT | 01-OCT-2004 (TrEMBLrel. 28, Last annotation update) | | | |
| DE | Mimbrain kinase (Hypothetical protein T04C10.1). | | | |
| GN | Name=mbk-1; Synonyms=704c10.1; | | | |
| OS | Caenorhabditis elegans. | | | |
| OC | Eukaryotes; Metazoa; Nematoda; Chromadorea; Rhabditota | | | |
| OC | Rhabditidae; Pelodertinae; Caenorhabditis. | | | |
| OX | NCBI_TaxID=6239; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=22505392; PubMed=12618396; | | | |
| RA | Raich W.B., Moorman C., Laceyfield C.O., Lehrer J., Bartsch D., | | | |
| RA | Plasterik R.H., Kandel E.R., Hobert O.; | | | |
| RT | "Characterization of Caenorhabditis elegans homologs of the Down | | | |
| RT | syndrome candidate gene DYRK1A."; | | | |
| RL | Genetics 163:571-580(2003). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=Br1stc1 N2; | | | |
| RX | MEDLINE=99069613; PubMed=9851916; | | | |
| RA | none; | | | |
| RT | "Genome sequence of the nematode C.elegans: A platform for | | | |
| RT | investigating biology."; | | | |
| RL | Science 282:2012-2018(1998). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=Br1stc1 N2; | | | |
| RA | Burton J.; | | | |
| RL | Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases. | | | |
| CC | -I- SIMILARITY: Belongs to the Ser/Thr protein kinase family. | | | |
| DR | EMBL; AY064464; AAL40874.1; -. | | | |
| DR | EMBL; Z69885; CAA93756.2; -. | | | |
| DR | HSSP; P24941; 1H0V. | | | |
| DR | Inact; Q8MQL7; -. | | | |
| DR | GO; GO:0005524; F:ATP binding; IEA. | | | |
| DR | GO; GO:0004674; F:protein serine/threonine kinase activity; IEA | | | |
| DR | GO; GO:0016740; F:transferase activity; IEA. | | | |
| DR | GO; GO:000468; F:protein amino acid phosphorylation; IEA. | | | |

DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser Thr kinase.
DR InterPro: IPR008271; Ser Thr_kin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00111; PROTEIN KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 882 AA; 101244 MW; D5C9DD354DA709 CRC64;

Query Match 11.3%; Score 248.5; DB 2; Length 882;
Best Local Similarity 25.7%; Pred. No. 4e-10;
Matches 104; Conservative 57; Mismatches 113; Indels 131; Gaps 17;
QY 88 NIVTLGVFTIHPSPVNSRCLLLELDVSVSELLYSHQSGMMWIGHCARDVLEALA 147
DB 389 NIVTLKGHP-VHRA-----HCLIVPELLSYLVYDLKNTSRGVSLNARKFAQGLKTL 443
QY 148 FLHHE--GVVHADLKERNILW-SAENECFRLIDFGLSPKGNODVXYIQTGYRADEAL 204
DB 444 FLSSPELSTIHCDELKRENVLLVNAKRSQIRVIDFGSSCGTHRYQYIGRFYRSEVLL 503
QY 205 QNCLAQGLQSPRECSAVDMLGILLLEMFSG-----MLQ----- 241
DB 504 -----GLAYDYK-----IDMWSLGLVEMHAGEPLFAGSSVDQMKEVLAGMPKKE 552
QY 242 -----KHTVRSQMKNSAIIIDHIFASRAVNAAI 272
DB 553 MLDIGKRTKHYPRKTEIDGIYCKKTDGYRHT-----YRARGAKLHEILG-----VTSGG 603
QY 273 PA-----YHLRDLIKSMLEDDPSRRIPEAKLCSPPFSI--PFAPIEDLV 316
DB 604 PGRRLGEPGSHVEDYSKFDLIKRMLOFDPKORI-----SPYYVVRHPLKQKEERY 656
QY 317 MLPTPLRLNLVDDDLVGNBEEVEDVEYKECKY--GPRVSLLVPRNGRQGVY 374
DB 657 PGPFPVSH-----SNLQOQOOLYIQOPSSQMSQVNESPSVSYY 696
QY 375 E-----YANAGDSKAAQKLTGRMFDGKF--VVAIFYPLSAKKGY 413
DB 697 EDNGMTRQAPGSSANPISVTSSFDGDADEVDAGRRRFSAHQNY 741

RESULT 9

Q6GM90 PRELIMINARY; PRT; 395 AA.
AC 06GM90;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haile F.,
RA Diatchenko L., Marinsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Strausberg R.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: BC074183; AH74183.1; -
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser Thr kinase.
DR InterPro: IPR008271; Ser Thr_kin_AS.
DR InterPro: IPR001245; Tyr kinase.
KW Pfam: PF00069; Pkinase; 1.
KW ProDom: PD000001; Prot kinase; 1.
KW SMART: SM00220; S_TKc; 1.
KW SMART: SM00219; TyrKc; 1.
KW PROSITE: PS00107; PROTEIN KINASE ATP; 1.
KW PROSITE: PS00111; PROTEIN KINASE_DOM; 1.
KW PROSITE: PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 395 AA; 43896 MW; 4C78BBF828A2CC4 CRC64;
Query Match 11.0%; Score 243.5; DB 2; Length 395;
Best Local Similarity 27.7%; Pred. No. 3.5e-10;
Matches 91; Conservative 47; Mismatches 108; Indels 83; Gaps 16;
QY 4 SGCAMGAEPPRFLFAFGRLMQVQSRIGSGSSASV-----YRVCCGNPGSPGA 52
DB 9 SGACMKKQ-----VEDIKRMFQKEVLTGAFSEVVLAEKERTGKLPVAVCI-----PKKA 59
QY 53 LKQFLPFGTTGAAGAAGYGFPRKEPAALEQLOGHNRIVTLVGVFTIHPSPVNSRCLL- 111
DB 60 LK-----GKXSSINEIEIAVLRIKIR-HENIVALEIDY-----ESPSSLYLW 99
QY 112 -----ELIDVSVSELLYSHQSGMMWIGHCARDVLEALFLHBEGVHADLKPRNTL 165
DB 100 QLVSGGELPD-RIVEKGFYTERKDASTL-----IQVLDVAVSYLRIGLVHDDLKRENLL 152
QY 166 WSAENECFRLI--DFGLSPKGNODVXYIQ--TDGYRAPEALQNCQAQGLQSPTECTS 221
DB 153 YFSQVBESEKIMISDFGLSGMEKGDVMSYAGTGPVVAPEV-----LAQ-----KPYSK 201
QY 222 AVDLMSLGILLLEMFSGMKLKTATVSQBEKANKNSAIIIDHIFASKAVNAAPAY-----H 276
DB 202 AVDCWSIGIYAILLTCGY-----PPFYDENDSRLFEQIL--KADYEPDPSYWDISBS 252
QY 277 LRDLIKSMLEDDPSRRIPEAKLCSPPFS 305
DB 253 ADFQIONLEKQDPNKKYTCQALRHFWIA 281

RESULT 10

09CA22
AC 09CA22 PRELIMINARY; PRT; 720 AA.
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Hypothetical protein T3288.5.
GN Name=T3288.5;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Bentto M., Creasy T.H., Haas B.J., Wu D.,
RA Maizl R., Rønning C.M., Koo H., Fujii C.Y., Uteback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nieman W.C., Fraser C.M.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AC012193; AAC51619.1; -.
DR PIR, A96807; A96807.
DR GO, GO:0005524; F:ATP binding; IEA.
DR GO, GO:0004674; F:Protein serine/threonine kinase activity; IEA.
DR GO, GO:0016740; F:transferase activity; IEA.
DR GO, GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro, IPR011009; Kinase_like.
DR InterPro, IPR000719; Prot_kinase.
DR InterPro, IPR002290; Ser_thr_kinase.
DR InterPro, IPR008271; Ser_thr_pkin_AS.
DR Pfam, PF00069; Pkinase; I.
DR Prodom, PD000001; Prot_kinase; 1.
DR SMART, SM00220; S_TKC; 1.
DR PROSITE, PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE, PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE, PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
DR ATP-binding; Hypothetical protein; Kinase; Transferase.
KW SEQUENCE 720 AA; 80027 MW; F26A9A0B1C6D91DE CRC64;
SQ

Query Match 11.0%; Score 243.5; DB 2; Length 720;
Best Local Similarity 25.9%; Pred. No. 7.5e-10;
Matches 106; Conservative 56; Mismatches 146; Indels 101; Gaps 17;
11 EPPRPLFAFGRLMOVQSRGSSASVYRVCCGPPGALAKOFLPPTGAASAAE 70
DB 331 DPULFRKVNKLYQRLGKISSGSSSEVHKV---SSDCTIYALKTKLGRDYATA--- 383
QY 71 YGFRKRAALEQIQGHRIIVTLVGVFTIHPSPNVPSCRLLLELDVSVSE----- 120
DB 384 YGFCQELGYKKLKGKTNITQL-----IDYEVTDKTLLEVLNGTMSKDGKVEDGF 436
QY 121 ---LLIYS---SHQCSMW-----MICHGARDVLEALFLHHEGYADLK 160
DB 437 IYVWLVYGEIDLAMHSQKREIEGSDRTIDENWMLRFYMOQIIQAVNTIHEERIVASDK 496
QY 161 PRRIILSAENECFKLIDFGSPKGNQDVKYIOTD-----GYRAPEALONCLAQAGQ 214
DB 497 PANFL--LVKGFKLIDFGIA-KAINSDTNIIQDSQVGLTWSMSEPAFCNSEDENG-- 551
QY 215 SDTECTSAVDLWSLGIILLEMFGSKLKTIVRSQEWKANSATIID--HIFASRAVYNAI 272
DB 552 NTKICGRSPDIWSLGLIYQWVG-RTRPADYKTFW-AKFKVITDPNHETTYQLSNP-- 607
QY 273 PAYHLDLISMLHDDPSRIIPAEMLCSPFSIPPAFH----- 311
DB 608 ---WLIDLMKCKCLAWRNQRMRIPELLQHPFLAPFI-PHEPQVKTIKLPSLIAESGSD 663
QY 312 -----IEDLVMLPTVLRLLNVLDLDDYLGNEEEYEDVADVEEC 351
DB 664 DKANSISQLEQLSNPAFLPR-----NDVLSRDQNOQLSRVSEL 706

RESULT 11
ID 025758 PRELIMINARY; PRT; 765 AA.
AC 025758
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Mitogen-activated protein kinase-related protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=9833;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=DD2;
RC lin D.T., Goldman N.D., Syn C.;
RT "Stage-specific expression of a Plasmodium falciparum protein related
to the eukaryotic mitogen-activated protein kinases.";
RL Mol. Biochem. Parasitol. 78:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=DD2;
RC lin D.T., Goldman N.D., Syn C.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL, U06377; AAC47170.1; -.
DR HSSP, Q16539; 1KVI.
DR GO, GO:0005524; F:ATP binding; IEA.
DR GO, GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO, GO:0016740; F:transferase activity; IEA.
DR GO, GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro, IPR011009; Kinase_like.
DR InterPro, IPR003527; MAP_kin.
DR InterPro, IPR000719; Prot_kinase.
DR InterPro, IPR002290; Ser_thr_kinase.
DR InterPro, IPR008271; Ser_thr_pkin_AS.
DR Pfam, PF00069; Pkinase; 1.
DR Prodom, PD000001; Prot_kinase; 1.
DR SMART, SM00220; S_TKC; 1.
DR PROSITE, PS01351; MARK; UNKNOWN_1.
DR PROSITE, PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE, PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE, PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 765 AA; 90231 MW; 5BEC8E6A4B02D32 CRC64;
SQ

Query Match 11.0%; Score 243.5; DB 2; Length 765;
Best Local Similarity 25.8%; Pred. No. 8.1e-10;
Matches 105; Conservative 58; Mismatches 145; Indels 99; Gaps 20;
23 WQVQSRIGSSGSASVYRVCCGPPGALAKOFLPPTGAASAA--AEYGRKRAAL 80
DB 23 YDLKKVKGAGVGFKGGCKKNKIV--AVAKIP-----GAFQCTDQARFR-EIIFL 74
QY 81 EQLQGHRIIVTLVGVFTIHPSPNVPSCRLLLELDVSVSEL---LIISSHQCSMMMIQ 136
DB 75 YELNGHDNIITKLMDYK--AKNDNDIYLIPEFMETDLEVIKADLLEIHK-----K 124
QY 137 HCARDTLEALFLHHEGYADLKPRNIIAMSLENK-FKLIDFGS-----FKEGNDV 189
DB 125 YIIYQLRLKTYIHGGGLHRDIKPSNII--VNSCHIKVADFGIARISITHVANKVPI 182
QY 190 ---KYIQTDGARAPEALQNCIAQAGLQSPTECTSAVDLWSLGIILLEMFGSKL---KHT 244
DB 183 LVDYATRYARAPFI-----LGGTHTEDEVDMMSLGCINGELLCKGRLPTGNSI 232
QY 245 VRSQE-----WKANSATIID--HIFASRAVYNAI-PAIPYHLR-----DLISK 284
DB 233 KMQLEKIIQIVGKPKMKDIEDIRSPFAEKTISSFVDLKKKNIKDIQYASNESLILLEKL 292
QY 285 LDDPSRRIPAEML-----CSPFSIPPAFHIEDVLMPTVLRLLNVLDLDDYLGNEERY 340

DB 293 LQFNPSKRISAENALKHKVEEFHSIIIDFPTCHIIITP----- 311
 QY 341 EDVVEVKECCQYGPVSLVLPKPNRGQVFEVYANMGDSKRAQK 387
 DB 332 --INDNTKRVNFYRVNVYFVIMRNKFKHSNVL-----NQGSEKKEK 372

RESULT 12

DB 084VX4 PRELIMINARY; PRT; 777 AA.
 AC 084VX4;
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE A16177720
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RX [1]
 RA SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
 RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Huan V.W.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
 RA Palm C.J., Quach H.T., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Tortum M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
 RA Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BR004636; AAC042882.1; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Prot. kinase.
 DR InterPro; IPR002290; Ser. Thr. kinase.
 DR InterPro; IPR008271; Ser. Thr. pkin. AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot. kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
 KW ATP-binding; Kinase; Transferase.
 SQ SEQUENCE 777 AA; 86376 MW; F2C25DA9609A878F CRC64;

Query Match 11.0%; Score 243.5; DB 2; Length 777;
 Best Local Similarity 25.9%; Pred. No. 8.2e-10;
 Matches 106; Conservative 56; Mismatches 146; Indels 101; Gaps 17;

QY 11 EPPRFLEAFRLMVQVSRLLGSSASAYVRCCGNPSPGALKQFLPPTGTAASAAE 70
 DB 388 DPLLPKVNKGKLYQRLEKIGSGSSSEVHKYI--SSDCTYALKKIKLKGROYATA---- 440
 QY 71 YGRKRRALAEQOGHRNIVLYGVFTIHSPVPSRCLLELDVSVS----- 120
 DB 441 YGCGEIGYIKKIKKGNIIQL-----IDYEVTIKTLLEQEVANGSMKSGRYKEDGF 493
 QY 121 ---LLYS-----SHQCGSMW-----MIOHCADVLEALAFLLHGGYVHADLK 160
 DB 494 IYVAVLEGEIDLAMLSQKKREIEGSRITIDENWLRVYQOIIQAVVTHIEEIVHSDLK 553
 QY 161 PNNILMSAENECFLIDFGLSFKREGNODVYIQT-----GYRAEALQNCIAQAGIQ 214
 DB 554 PANFL--LVNGFLKLIDFGIA-KAINSDTNIQRDSQVGLTSMSPFAFVCSNDENG-- 608
 QY 215 SDRECHSAVDNLGILILEMFGMKLKTIVRSQEKANSALID--HIFASAVVNAAL 272
 DB 609 NTKICRPSDWSWIGCLLYQVYVG-RTPPADYKTFW-AKPKVITDPNHEITYNQLSNP-- 664
 QY 273 PAYHLMDLKSMLHDDPSRRIPIAMMALCSPFFSIFPAFH----- 311

DB 665 ---WLIDLMKKCLANDRNRWRIPBLQHPFLAPPI-PHEPQVTKIKLPSLIAESGSD 720
 QY 312 -----IEDLVMLPTFVLRLNLVDDDYIGNEBEEDVVEDYKEEC 351
 DB 721 DRANSMISQLEQLLSNPAPLPR-----NDVIDSRDQNOQLSRVSELC 763

RESULT 13

DB 094656 PRELIMINARY; PRT; 826 AA.
 AC 094656;
 DT 01-FEB-1997 (TReMBLrel. 02, Created)
 DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Mitogen-activated protein kinase 1, serine/threonine protein
 DE kinase.
 GN Name=pfmap1;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 NCBI_TaxID=5833;
 RX [1]
 RA SEQUENCE FROM N.A.
 RA STRAIN=IT04;
 RX MEDLINE=97080491; PubMed=8921836;
 RA Doering C.D., Parzy D., Langelley G., Horrocks P., Carter R.,
 RA "A map kinase homologue from the human malaria parasite Plasmodium
 RT falciparum.";
 RL Gene 177:1-6(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IT04;
 RA Doering C.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- Similarity: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; X82646; CAA57972.1; -.
 DR PIR; JC5153; JC5153.
 DR HSSP; Q16539; 1KV1..
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR003527; MAP kin.
 DR InterPro; IPR000719; Prot. kinase.
 DR InterPro; IPR002290; Ser. Thr. kinase.
 DR InterPro; IPR008271; Ser. Thr. pkin. AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS01351; MARK; UNKNOWN_1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 826 AA; 96645 MW; A668EF77E73438A CRC64;

Query Match 11.0%; Score 243.5; DB 2; Length 826;
 Best Local Similarity 25.8%; Pred. No. 8.9e-10;
 Matches 105; Conservative 58; Mismatches 145; Indels 99; Gaps 20;

QY 23 NOVQSRLLGSSASAYVRCCGNPSPGALQFLPPTGTAASAA--AEYFRKERAL 80
 DB 23 YDILKKVKGAGVGVFKGCKKNKIV--AVKKIF-----GAFQCTDQRTFR-EIIFL 74
 QY 81 EOLQGRNIVTVLYGVFTIHSPVPSRCLLELDVSVSEL-----LYSSHQCGSMMLIQ 136
 DB 75 YELNGHNDIITKMDYK--AKNDNDIYIIFPMTDLHVIKADLBEIHK-----K 124
 QY 137 HCARVDLEALFLHGGYVHADLKPRNIIISANEC-FKLIDFGLS-----FKEGNODV 189
 DB 125 YIIYQLRLAKYTHSGGLLRDIPKSNIL--VNSECHIKVADFGIARISISTVHNENKVP 182

QY 190 --KYIOTDGYRABEALONCLAQIQSPTCECTSAVDLSLGIILLEMESGML---KKT 244
 Db 183 LNDYATRYRABEI-----LLGSTHYTEDVDMMSLGGCIMGELLCGKPLFTGNST 232
 QY 245 VNSQE-----WKANSSAIID--HIPASKAVVN-AAIPAYHNR-----DLIKSM 284
 Db 233 MNQLEKIIQVIGKPNKKDIEDIRSPFAEKIISFVDLKKKNLKDICYKASNESLIDLEK 292
 QY 285 LHDDPSRRIPAEAL-----CSPFESIPEAPHIEDLVMLPTPVLRLLNLVDDYLGNEEY 340
 Db 293 LQNPSPKRSIAENALKHKYVEEFHSIIDEPTRHIIITP----- 331
 QY 341 EDVEDVEKEECQKYGVPVSLVPEKNPGRGQVFVEYANAGSKAOK 387
 Db 332 --INDNTKRVNFYRNVVYFVIMRRNKKFHSNVL-----NGESKKEEK 372

RESULT 14
 Q8ILFO PRELIMINARY; PRT; 914 AA.

AC 08ILFO;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Mitogen-activated protein kinase 1.
 GN ORFNames=PF14_0294;
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=63329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22255705; PubMed=12368864;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlson J.M., Hall N., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyte S.,
 RA Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
 RA Pertea M., Allen J., Selengut J., Haft D., Mather W.W., Valdivia A.B.,
 RA Martin D.M., Fairclamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrrell B.,
 RT "Genome sequence of the human malaria parasite Plasmodium falciparum."
 RL Nature 419:498-511(2002).
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AB014820; AAN36907.1; -.
 DR HSSP; Q16539; 1KVL.
 DR GO; GO:000524; P:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; E:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR003527; MAP kin.
 DR InterPro; IPR000719; Ser kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_kinase.
 DR Pfam; PF00069; Kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS01351; MAPK; UNKNOWN 1.
 DR PROSITE; PS01017; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 914 AA; 107277 MW; CAAP486AB051FA72 CRC64;

Query Match 11.0%; Score 243.5; DB 2; Length 914;
 Best Local Similarity 25.8%; Pred. No. 1e-09;
 Matches 105; Conservative 58; Mismatches 145; Indels 99; Gaps 20;

QY 23 MOVOSRLSGSGSASVYRVRCCGNPGSPGALQKFLPPTGTAASA--AEYGRKRRAL 80
 Db 23 YDLKKVGGKAYGVVFRGCRCKKNIV--AVKKIF-----GAPQNCIDQQRTR-ETIFL 74

QY 81 EQLQGRNIVTLVGYFTTHFSPNVSRCILLELDVSVSEL---LTVSSHQCSMMMIQ 136
 Db 75 YELNGHDNIITKLMDVYK---AKNDNDIYLIFDPFMTDLHEVIKADLEIRK-----K 124
 QY 137 HCARVLEALFLHHEGVYHADLKPRLMSAENSC-FLLIDFGLS-----FKENQDV 189
 Db 125 YIIYQLRLKTIKTHSGGLHARDIKPSNLL--VNSCHIKVADFGIARISITVNEKVP 182
 QY 190 --KYIOTDGYRABEALONCLAQIQSPTCECTSAVDLSLGIILLEMESGML---KKT 244
 Db 183 LNDYATRYRABEI-----LLGSTHYTEDVDMMSLGGCIMGELLCGKPLFTGNST 232
 QY 245 VNSQE-----WKANSSAIID--HIPASKAVVN-AAIPAYHNR-----DLIKSM 284
 Db 293 LQNPSPKRSIAENALKHKYVEEFHSIIDEPTRHIIITP----- 331
 QY 341 EDVEDVEKEECQKYGVPVSLVPEKNPGRGQVFVEYANAGSKAOK 387
 Db 332 --INDNTKRVNFYRNVVYFVIMRRNKKFHSNVL-----NGESKKEEK 372

RESULT 15
 Q8AYR3 PRELIMINARY; PRT; 395 AA.

AC 08AYR3;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Calcium/calmodulin-dependent protein kinase (Cam-KI protein).
 GN Name=Cam-KI;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Saneyoshi T., Kume S., Mikoehiba K.,
 RT "Calcium/calmodulin-dependent protein kinase I in Xenopus laevis."
 RL Comp. Biochem. Physiol. B, Comp. Biochem. 134:499-507(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocytes;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.U., Malek U.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywicki M.I., Skalska U., Smalios D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocytes;
 RX MEDLINE=22341132; PubMed=12454917;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
RN (4)
RP SEQUENCE FROM N.A.
RC TISSUE=Oocytes;
RA Klein S., Strausberg R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB082399; EAC13847.1; -;
DR EMBL; BC070745; AAH70745.1; -;
DR HSSP; Q63450; 1A06.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; F:Protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:000468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 395 AA; 43801 MW; 5381D04890B5D3AE CRC64;

Query Match 11.0%; Score 242.5; DB 2; Length 395;
Best Local Similarity 27.7%; Pred. No. 4.2e-10;
Matches 91; Conservative 47; Mismatches 108; Indels 83; Gaps 16;

QY 4 SGCWGAEPPEPRFLFAFGRLMQVQSRIGSGSSASV-----YVRCCGNGSPGPA 52
DB 9 SGTGCKKQ---VEDIKRMQFKFVGTGAFSEVLAEEKATGKLPVVKCI-----PKKA 59
QY 53 LKQFLPFGTTGAASAAYGFRKERAALBQLOGRNIVTLVGVFTIHFSNVPSRCLL- 111
DB 60 LK-----GKESIENEIVLRKIK-HENIVALEDIY-----ESPSHLYLVA 99
QY 112 -----ELDVSVSELLVSSHQCSMMMIQHCARDVLEALAFHHEGYVHADLKPRNII 165
DB 100 QLVSGSELFD-RIVEKGFTYKXASTL-----IKQVLDAVSYLHRLGIVHRDLKPRNII 152
QY 166 WSAENECEFKLI--DFGLSFKEGNQDYKVIQ--TDGYRAPAEILQNCIAQAGLSDTCTG 221
DB 153 YFSQVESKIMISDFGLSKMEGKGDVWSTACGPGYVAPV-----LAQ-----KPYSK 201
QY 222 AYDLMLSLGIIILEMFGMLKHTVRQEWKANSALIIDHIFASKAVVNAALPAY-----H 276
DB 202 AYDCWSIGVAYLILCGY-----PPFYDENDSKLFEQTL--KADYEFDSFYWDIDSDS 252
QY 277 LRLDIKSMLDPPSRRIAPAMALCSPFFS 305
DB 253 AKDFIQNLMEKDPNKRYPACEQALRHPRIA 281

Search completed: November 29, 2004, 11:46:52
Job time : 126.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 109.5 Seconds
(without alignments)
1372.672 Million cell updates/sec

Title: US-10-798-532-2

Perfect score: 2206
Sequence: 1 MAGSCAMAEPPRFLIARG.....VATFVPLSAVKKGYLQTL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1 | 2198 | 99.6 | 419 | 4 | AAG65767 |
| 2 | 2198 | 99.6 | 419 | 5 | ABP64797 |
| 3 | 2191 | 99.3 | 419 | 4 | AAU03534 |
| 4 | 2191 | 99.3 | 419 | 4 | AAE11769 |
| 5 | 2190.5 | 99.3 | 418 | 3 | AAV70305 |
| 6 | 2187.5 | 99.2 | 418 | 3 | AAV70306 |
| 7 | 2186 | 99.1 | 419 | 7 | ADDD48093 |
| 8 | 483.5 | 21.9 | 137 | 5 | ABP64662 |
| 9 | 243.5 | 11.0 | 638 | 3 | AAG38502 |
| 10 | 243.5 | 11.0 | 720 | 3 | AAG38501 |
| 11 | 243.5 | 11.0 | 777 | 3 | AAG38500 |
| 12 | 239.5 | 10.9 | 544 | 2 | AAW85028 |
| 13 | 238 | 10.8 | 298 | 7 | ADDD21398 |
| 14 | 236 | 10.7 | 435 | 6 | AAE31916 |
| 15 | 235 | 10.7 | 367 | 3 | AAE21317 |
| 16 | 235 | 10.7 | 435 | 6 | AAE31917 |
| 17 | 235 | 10.7 | 435 | 6 | AAE31922 |
| 18 | 235 | 10.7 | 435 | 6 | AAE31921 |
| 19 | 235 | 10.7 | 435 | 6 | AAE31914 |
| 20 | 235 | 10.7 | 435 | 6 | AAE31913 |
| 21 | 235 | 10.7 | 544 | 2 | AAW85029 |
| 22 | 234.5 | 10.6 | 433 | 3 | AAE19094 |
| 23 | 234 | 10.6 | 433 | 8 | ADP87472 |
| 24 | 234 | 10.6 | 298 | 2 | AAW85689 |
| 25 | 234 | 10.6 | 298 | 7 | ADDD21402 |

ALIGNMENTS

| | | | | | | | |
|--|----|-------|------|-----|---|-----------|---------------------|
| RESULT 1 | 26 | 234 | 10.6 | 298 | 7 | ADDD21406 | AdD21406 Rat cycli |
| AA65767 | 27 | 234 | 10.6 | 298 | 2 | ADDD21394 | AdD21394 Mouse cyc |
| AA65767 standard, protein, 419 AA. | 28 | 234 | 10.6 | 416 | 7 | AAW06795 | AAW06795 High Osmo |
| AA65767; | 29 | 234 | 10.6 | 435 | 6 | AAE31910 | AAE31910 Yeast Hog |
| 07-JUN-2002 (first entry) | 30 | 234 | 10.6 | 435 | 6 | AAE31915 | AAE31915 Yeast Hog |
| Human protein kinase polypeptide 13887. | 31 | 234 | 10.6 | 435 | 6 | AAE31918 | AAE31918 Yeast Hog |
| Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human; cytostatic; antiinflammatory; immunosuppressive; cardiac; hepatocytic; Gene therapy. | 32 | 234 | 10.6 | 435 | 6 | AAE31919 | AAE31919 Yeast Hog |
| Homo sapiens. | 33 | 234 | 10.6 | 435 | 6 | AAE31926 | AAE31926 Yeast Hog |
| WO200173050-A2. | 34 | 234 | 10.6 | 435 | 6 | AAE31911 | AAE31911 Yeast Hog |
| 04-OCT-2001. | 35 | 234 | 10.6 | 435 | 6 | AAE31909 | AAE31909 Yeast Hog |
| 23-MAR-2001; 2001WO-US009483. | 36 | 234 | 10.6 | 435 | 6 | AAE31912 | AAE31912 Yeast Hog |
| 24-MAR-2000; 2000US-0191846P. | 37 | 234 | 10.6 | 435 | 6 | AAE31927 | AAE31927 Yeast Hog |
| (MILL-) MILLENNIUM PHARM INC. | 38 | 233 | 10.6 | 435 | 6 | ABR58686 | ABR58686 Human can |
| Meyers R; | 39 | 233 | 10.6 | 298 | 7 | ADDD21410 | ADDD21410 Goldfish |
| WPI; 2001-611632/70. | 40 | 233 | 10.6 | 298 | 7 | ADDD21392 | ADDD21392 Human cyc |
| N-PSDB; AA166828, AA166829. | 41 | 233 | 10.6 | 298 | 7 | ADP45038 | ADP45038 Human kin |
| New human protein kinase polypeptides, 3714, 16742, 23546 and 13887, useful in diagnosis of cancer or cellular proliferation or to screen for polypeptide modulators useful to treat such conditions. | 42 | 233 | 10.6 | 298 | 8 | AD157232 | AD157232 NK1AWE d |
| Claim 4; Fig 10A-C; 16pp; English. | 43 | 233 | 10.6 | 298 | 8 | ADJ66599 | ADJ66599 cdk2 prot |
| The invention provides novel human protein kinase polypeptides, 3714, 16742, 23546 and 13887 and nucleic acid molecules encoding them. The protein kinase polypeptides can be expressed by standard recombinant methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides are useful for diagnostic and screening methods to identify subjects (at risk of) having cancer or cellular proliferation and/or differentiation disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and modulators are useful for the treatment of cancer, particularly colon cancer or cellular proliferation and/or differentiation disorders. Other disorders associated with 3714, 16742, 23546 or 13887 expression or | 44 | 233 | 10.6 | 392 | 4 | ABB63118 | ABB63118 Drosophila |
| | 45 | 232.5 | 10.5 | 355 | 4 | AAE50055 | AAE50055 Murine De |

CC activity that can be treated include bone related disorders, inflammatory
CC disorders, autoimmune diseases, cardiovascular disorders and liver
CC diseases. The present sequence represents a human protein kinase
CC polypeptide 13887

XX Sequence 419 AA;

Query Match 99.6%; Score 2198; DB 4; Length 419;
Best Local Similarity 99.8%; Pred. No. 2,1e-229;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGSGCAGMBPRLPFLAEGRLMVOVSRLGSSGSASVYRVRCGNGSPGALKQFLPPG 60
DB 1 MAGSGCAGMBPRLPFLAEGRLMVOVSRLGSSGSASVYRVRCGNGSPGALKQFLPPG 60
QY 61 TTGAASAAEYGFRRKRALAEQLQGHNRNVTLYGVFTIHFSNVPSRCLLELDVSVSE 120
DB 61 TTGAASAAEYGFRRKRALAEQLQGHNRNVTLYGVFTIHFSNVPSRCLLELDVSVSE 120
QY 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
DB 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
QY 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
DB 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
QY 181 SFKEGQDVKYIOTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIIILEMFSGMK 240
DB 181 SFKEGQDVKYIOTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIIILEMFSGMK 240
QY 241 LKHTVRSQEWKANSIIIDHIFASKAVVNAIPAYHLDIKSMHDDPSRIIPAEMLC 300
DB 241 LKHTVRSQEWKANSIIIDHIFASKAVVNAIPAYHLDIKSMHDDPSRIIPAEMLC 300
QY 301 SPFFSIPFAPHIEDLVMLPTPVLRLNVLDLDDYLGNEBEYEDVDEVECKQKGPVSVL 360
DB 301 SPFFSIPFAPHIEDLVMLPTPVLRLNVLDLDDYLGNEBEYEDVDEVECKQKGPVSVL 360
QY 361 LVPKENPGRGVVEVYANAGDSKAAQKLLTGMFPGKFVATFYPLSAVKRGYLYQTLL 419
DB 361 LVPKENPGRGVVEVYANAGDSKAAQKLLTGMFPGKFVATFYPLSAVKRGYLYQTLL 419

RESULT 2
ABP64797
ID ABP64797 standard; protein; 419 AA.

XX ABP64797;

XX 25-FEB-2003 (first entry)

DE Human protein SEQ ID 457.

XX Human; expressed sequence tag; EST; hematopoietic disorder;
KW central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; noctropic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective.

XX Homo sapiens.

XX WO200259260-A2.

XX 01-AUG-2002.

XX 16-NOV-2001; 2001MO-US042950.

XX 17-NOV-2000; 2000US-00714936.

XX (HYSB-) HYSBQ INC.

XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI; 2002-590824/63.
DR N-PSDB; ABQ9383.

PT New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity.

PS Claim 20; SEQ ID NO 457; 394bp; English.

CC The present invention relates to novel human coding sequences (ABQ99268-
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC therapeutic, diagnostic and research methods. The polynucleotides may be
CC used in the field of molecular biology as hybridisation probes, primers
CC for PCR, for chromosome and gene mapping, for the recombinant production
CC of protein, or in generation of anti-sense DNA or RNA. The
CC polynucleotides are useful in diagnostics as expressed sequence tags
CC (ESTs) for identifying expressed genes or for physical mapping of the
CC human genome. The proteins may be used as molecular weight markers, or as
CC nutritional sources or supplements. The proteins may be used to maintain
CC and expand cell population in a totipotent or pluripotent state
CC useful for re-engineering damaged or diseased tissues, transplantation,
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
CC polynucleotides and proteins are useful for preventing, treating or
CC ameliorating disorders involving aberrant protein expression or
CC biological activity, e.g. hematopoietic disorders, central/peripheral
CC nervous system diseases, mechanical and traumatic disorders, non-healing
CC wounds, immune deficiencies and disorders, infectious diseases caused by
CC viral, bacterial or fungal infection, autoimmune disorders, allergic
CC reactions and conditions, coagulation disorders, or cancer. The
CC polynucleotide sequences of the invention were assembled from ESTs
CC isolated mainly by sequencing by hybridisation, and in some cases,
CC sequences obtained from one or more public databases. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPRO at
CC ftp.wipro.int/publicated_pct_sequences

XX Sequence 419 AA;

Query Match 99.6%; Score 2198; DB 5; Length 419;
Best Local Similarity 99.8%; Pred. No. 2,1e-229;
Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGSGCAGMBPRLPFLAEGRLMVOVSRLGSSGSASVYRVRCGNGSPGALKQFLPPG 60
DB 1 MAGSGCAGMBPRLPFLAEGRLMVOVSRLGSSGSASVYRVRCGNGSPGALKQFLPPG 60
QY 61 TTGAASAAEYGFRRKRALAEQLQGHNRNVTLYGVFTIHFSNVPSRCLLELDVSVSE 120
DB 61 TTGAASAAEYGFRRKRALAEQLQGHNRNVTLYGVFTIHFSNVPSRCLLELDVSVSE 120
QY 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
DB 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
QY 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
DB 121 LLLYSSHQGCSMMWIOHCARDVLEALAFHHEGYHADLKPRNIIIMSANECEFKLIDFGL 180
QY 181 SFKEGQDVKYIOTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIIILEMFSGMK 240
DB 181 SFKEGQDVKYIOTDGYRAPEALQNCIAQAGLQSDTECTSAVDLMSLGIIILEMFSGMK 240
QY 241 LKHTVRSQEWKANSIIIDHIFASKAVVNAIPAYHLDIKSMHDDPSRIIPAEMLC 300
DB 241 LKHTVRSQEWKANSIIIDHIFASKAVVNAIPAYHLDIKSMHDDPSRIIPAEMLC 300
QY 301 SPFFSIPFAPHIEDLVMLPTPVLRLNVLDLDDYLGNEBEYEDVDEVECKQKGPVSVL 360
DB 301 SPFFSIPFAPHIEDLVMLPTPVLRLNVLDLDDYLGNEBEYEDVDEVECKQKGPVSVL 360
QY 361 LVPKENPGRGVVEVYANAGDSKAAQKLLTGMFPGKFVATFYPLSAVKRGYLYQTLL 419
DB 361 LVPKENPGRGVVEVYANAGDSKAAQKLLTGMFPGKFVATFYPLSAVKRGYLYQTLL 419

RESULT 3

AAU03534
 ID AAU03534 standard; protein; 419 AA.
 AC AAU03534;
 XX
 DT 12-SEP-2001 (first entry)
 DE Human protein kinase #34.
 XX
 KW Human; protein kinase; PK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder.
 OS Homo sapiens.
 PN MO200138503-A2.
 PD 31-MAY-2001.
 PF 22-NOV-2000; 2000MO-US032085.
 PR 24-NOV-1999; 99US-0167482P.
 XX
 XX (SUGB-) SUGEN INC.
 PA
 PI Piowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 XX
 DR WPI; 2001-343950/36.
 XX N-PSDB; AAS06734.
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections.
 XX
 PS Claim 7; Fig 2; 433pp; English.
 XX
 CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
 CC protein kinases have been identified as members of the tyrosine or
 CC serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of hematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be used for
 CC gene therapy, and as DNA probes in diagnostic assays. The protein kinase
 CC polypeptides may be used as antigens in the production of antibodies
 CC against the protein kinases and in assays to identify modulators of
 CC protein kinase expression and activity
 CC
 SQ Sequence 419 AA;
 Query Match 99.3%; Score 2191; DB 4; Length 419;
 Best Local Similarity 99.5%; Pred. No. 1,2e-228;
 Matches 417; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 181 SPKSGNDVYKTIQTDSYRAPEAEIOWCLAQAGLSPTECTSAVDLMSLGIILLBMPGSK 240
 DB 181 SPKSGNDVYKTIQTDSYRAPEAEIOWCLAQAGLSPTECTSAVDLMSLGIILLBMPGSK 240
 QY 241 LKHTVASQEMKANSNAIIDHIFASKAVVNAAPAYHLRLIKSMHDDPSRRIPAEMLC 300
 DB 241 LKHTVASQEMKANSNAIIDHIFASKAVVNAAPAYHLRLIKSMHDDPSRRIPAEMLC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDYLGNEEYEDVEDVEKCECKYGPVSL 360
 DB 301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDYLGNEEYEDVEDVEKCECKYGPVSL 360
 QY 361 LVPKGNPRGQVFEVYANAGDSKAOKLITGSMFDSKPVATFPYLSAYKRGVLTLL 419
 DB 361 LVPKGNPRGQVFEVYANAGDSKAOKLITGSMFDSKPVATFPYLSAYKRGVLTLL 419
 RESULT 4
 ID AAEL1769 standard; protein; 419 AA.
 AC AAEL1769;
 XX
 DT 18-DEC-2001 (first entry)
 DE Human kinase (PKIN) -3 protein.
 XX
 KW Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
 KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
 KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;
 KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
 KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
 KW asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
 KW antimicrobial; cardiovascular; antiinflammatory; vaccine.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 23..304
 FT /note="Eukaryotic protein kinase domain"
 FT 48..289
 FT /note="Protein kinase domain"
 FT 74..293
 FT /note="Protein kinase domain"
 FT 108..241
 FT /note="Protein kinase domain"
 FT 148..166
 FT /note="Protein kinase catalytic domain"
 FT 223..245
 FT /note="Tyrosine kinase catalytic domain"
 FT 223..241
 FT /label="Transmembrane_domain"
 FT 273..295
 FT /note="Tyrosine kinase catalytic domain"
 FT 305..344
 FT /note="Serine/threonine protein kinase"
 FT 319..404
 FT /note="Ribonucleoprotein repeat"
 FT 320..412
 FT /note="Splicing factor like protein"
 PN MO200181555-A2.
 PD 01-NOV-2001.
 PF 20-APR-2001; 2001MO-US012992.
 XX
 XX 20-APR-2000; 2000US-0199021P.
 PR 28-APR-2000; 2000US-0200226P.
 PR 05-MAY-2000; 2000US-0202339P.

PR 11-MAY-2000; 2000US-0203505P.
 PR 16-MAY-2000; 2000US-0205564P.
 PR 26-MAY-2000; 2000US-0207739P.
 PR 01-JUN-2000; 2000US-0208795P.
 XX
 XX
 PA (INCY-) INCYTE GENOMICS INC.

XX Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
 PI Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
 PI Patterson C, Burrill JD, Marcus GA, Zingler KA, Reardon SA, Lu Y;
 PI Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
 PI Walsh RT, Ramkumar J, Borowsky ML, Au-Young J, Hillman JJ,
 PI Gururajan R;
 XX
 XX
 DR WPI: 2001-611740/70.
 DR N-PSDB; AAD18818.

XX Human kinases and nucleic acids, useful for preventing diagnosing and
 PT treating cancers, inflammation and immune disorders.
 XX

PS Claim 1; Page 125-126; 16pp; English.

XX The present invention relates to human kinases (PKIN) and the nucleic
 CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
 CC used in the prevention, diagnosis and treatment of diseases cancers,
 CC adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
 CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
 CC gout, microbial infections, cardiovascular disease and/or inflammation,
 CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
 CC infarction, cataract, growth and development disorder, seizure disorder,
 CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
 CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
 CC PKIN may be used to treat disorders associated with decreased PKIN
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of PKIN by expressing inactive proteins or to
 CC supplement the patient's own production of PKIN. PKIN nucleic acids may be
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into
 CC a host cell and culturing the cell to express the protein. PKIN nucleic
 CC acids and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acid sequences in samples and therefore which patients may be in
 CC need of restorative therapy. The present sequence is human PKIN-3 protein
 CC
 XX

Sequence 419 AA;

Query Match 99.3%; Score 2191; DB 4; Length 419;

Best Local Similarity 99.5%; Pred. No. 1.2e-228; Mismatches 2; Indels 0; Gaps 0;

Matches 417; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGSGCAGCAEPPRFLAEFGRLMQVOSRLGSSGSSASVYVRCCGNGSPGALKQFLPPG 60
 DB 1 MAGSGCAGCAEPPRFLAEFGRLMQVOSRLGSSGSSASVYVRCCGNGSPGALKQFLPPG 60
 QY 61 TTGAASAABEYGRKERRALEOQGRNIVTLVGVFTTHSPNVRSCILLELDVSVSE 120
 DB 61 TTGAASAABEYGRKERRALEOQGRNIVTLVGVFTTHSPNVRSCILLELDVSVSE 120
 QY 121 LLLYSSHOGCSMMWIGHCQADVEALAFHLHBEYVHADLPRIILMSAENECKLIDFGL 180
 DB 121 LLLYSSHOGCSMMWIGHCQADVEALAFHLHBEYVHADLPRIILMSAENECKLIDFGL 180
 QY 181 SFKEGNDVYIQTDDYRAPEALQNCIAQAGLOSTECTSAVDLMSLGIILLEMFSGMK 240
 DB 181 SFKEGNDVYIQTDDYRAPEALQNCIAQAGLOSTECTSAVDLMSLGIILLEMFSGMK 240
 QY 241 LKHTVRSQEWKANSIIDIHIFASKAVVNAIPAHLRDLIKSMHDDPSRRIPAMALC 300
 DB 241 LKHTVRSQEWKANSIIDIHIFASKAVVNAIPAHLRDLIKSMHDDPSRRIPAMALC 300
 QY 301 SPFFSTIPAPHTIEDLWMLPTPVRLNLVLDLDDYLGNEEEDVVEVKECKQYGVVSL 360
 DB 301 SPFFSTIPAPHTIEDLWMLPTPVRLNLVLDLDDYLGNEEEDVVEVKECKQYGVVSL 360

QY 361 LVPKENPGRGQVFEVYANAGDSKAAQKLLTGRMFDGKFAVATFPLSAVKRGYLYQTLL 419
 DB 361 LVPKENPGRGQVFEVYANAGDSKAAQKLLTGRMFDGKFAVATFPLSAVKRGYLYQTLL 419

RESULT 5
 AAY70305
 ID AAY70305 standard; protein; 418 AA.

XX AAY70305;

DT 06-JUN-2000 (first entry)

DE Human wild type serine/threonine kinase KIS (hKIS) protein.

XX KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
 KW cyclin-dependent kinase inhibitor; CKI; p27; modulator; treatment;
 KW cell proliferative disease; vascular disorder; gene therapy; restenosis;
 KW atherosclerosis.
 XX

OS Homo sapiens.

EH Key Location/Qualifiers

FT Misc-difference 186 /note= "Encoded by AATCAG"

XX MO200011165-A1.

PD 02-MAR-2000.

PF 20-AUG-1999; 99WO-US018903.

PR 21-AUG-1998; 98US-0097710P.

XX (NABE/) NABEL G J.

PA (NABE/) NABEL E G.

PI Nabel GJ, Nabel EG;

DR WPI: 2000-237648/20.

DR N-PSDB; AA251355.

XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
 PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 XX proliferation.

PS Example 1; Page 57-58; 70pp; English.

XX The present protein sequence is the wild type human KIS (hKIS), isolated
 CC from a human B-cell library, using a yeast two hybrid screening system.
 CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
 CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS controls cell
 CC proliferation and is localised predominantly in the nucleus. It binds to
 CC CKI p27 and inhibits its ability to arrest cells in G1 phase. It has an
 CC amino terminal serine/threonine kinase consensus region and a carboxy
 CC terminus with 42% sequence similarity to hU2AF65, a subunit of the
 CC splicing factor U2AF. The hKIS sequences are used to modulate cell
 CC proliferation and treat cell proliferative and vascular diseases. The DNA
 CC sequence may be used in gene therapy to treat vascular disorders such as
 CC restenosis or atherosclerosis
 CC
 XX

Sequence 418 AA;

Query Match 99.3%; Score 2190.5; DB 3; Length 418;

Best Local Similarity 99.8%; Pred. No. 1.4e-228; Mismatches 0; Indels 1; Gaps 1;

Matches 418; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAGSGCAGCAEPPRFLAEFGRLMQVOSRLGSSGSSASVYVRCCGNGSPGALKQFLPPG 60
 DB 1 MAGSGCAGCAEPPRFLAEFGRLMQVOSRLGSSGSSASVYVRCCGNGSPGALKQFLPPG 60
 QY 61 TTGAASAABEYGRKERRALEOQGRNIVTLVGVFTTHSPNVRSCILLELDVSVSE 120
 DB 61 TTGAASAABEYGRKERRALEOQGRNIVTLVGVFTTHSPNVRSCILLELDVSVSE 120

DB 61 TTGAASAAYGFRKERALBQLQGHRIIVTLGVFTIHFSPNVPSCRLLLELDVSVSE 120
 QY 121 LLLYSHQCCSMMWIOHCARDVLEALAFIHHGYVHADLKPRNIIIMSAENECFKLIDFGL 180
 DB 121 LLLYSHQCCSMMWIOHCARDVLEALAFIHHGYVHADLKPRNIIIMSAENECFKLIDFGL 180
 QY 181 SPKEGNODVKYIQTODGYRAPEALONCLAQAGLSQDTECTSAVDLSLGIILLEMFSGMK 240
 DB 181 SPKEGN-DVKYIQTODGYRAPEALONCLAQAGLSQDTECTSAVDLSLGIILLEMFSGMK 239
 QY 241 LKHTVRSQEMKANSNAIIDIHIFASKAVVNAAIIPAYHLRDLIKSMHDDPSRRIIPAEALC 300
 DB 240 LKHTVRSQEMKANSNAIIDIHIFASKAVVNAAIIPAYHLRDLIKSMHDDPSRRIIPAEALC 299
 QY 301 SPFFSIPFAPHIEDLVMLPTPLRLNLVDDYLGNEBEYEDVVDVEECCQYGPVSVL 360
 DB 300 SPFFSIPFAPHIEDLVMLPTPLRLNLVDDYLGNEBEYEDVVDVEECCQYGPVSVL 359
 QY 361 LVPKENPGRGVFEVYANAGDSKAAOKLLTGRMFDGKPVVATFPLSAIKRGILYQTL 419
 DB 360 LVPKENPGRGVFEVYANAGDSKAAOKLLTGRMFDGKPVVATFPLSAIKRGILYQTL 418

RESULT 6
 AAY70306
 ID AAY70306 standard; protein; 418 AA.

AC AAY70306;

DT 06-JUN-2000 (first entry)

DE Human transdominant mutant serine/threonine kinase KIS (hKIS) protein.

KW KIS; human; hKIS; serine/threonine kinase; transdominant mutant; p27;
 KW G1 phase; cell proliferation; cyclin-dependent kinase inhibitor; CKI;
 KW modulator; treatment; cell proliferative disease; vascular disorder;
 KW gene therapy; atherosclerosis; restenosis.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 54

FT /note= "Wild type Lys substituted with Arg"

FT Misc-difference 186

FT /note= "Encoded by AATCAG"

PN WO20001165-A1.

PD 02-MAR-2000.

PF 20-AUG-1999; 99WO-US018903.

PR 21-AUG-1998; 98US-0097710P.

PA (NABE/) NABEL G J.

PA (NABE/) NABEL E G.

PI Nabel GJ, Nabel EG;

DR WPI; 2000-237648/20.

DR N-PSDB; AA251356.

PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
 for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 proliferation.

PS Example 1; Page 59-61; 70pp; English.

CC The present protein sequence is the transdominant mutant human KIS
 CC (hKIS), constructed by site directed mutagenesis. A single nucleotide
 CC substitution (A to G) results in a lysine to arginine change in the
 CC protein sequence. hKIS is a serine/threonine kinase, that acts as an
 CC inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS

CC controls cell proliferation and is localised predominantly in the
 CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
 CC G1 phase. The hKIS sequences are used to modulate cell proliferation and
 CC treat cell proliferative and vascular diseases. The polynucleotide
 CC sequence may be used in gene therapy to treat vascular disorders such as
 CC restenosis or atherosclerosis

SO Sequence 418 AA;

Query Match 99.2%; Score 2187.5; DB 3; Length 418;
 Best Local Similarity 99.5%; Pred. No. 2.9e-228;
 Matches 417; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 MAGSGCAGCAEPPRLLEAFGRIMQVSRGSSSAYVRCCGPGSPGALXQFLPPG 60
 DB 1 MAGSGCAGCAEPPRLLEAFGRIMQVSRGSSSAYVRCCGPGSPGALXQFLPPG 60
 QY 61 TTGAASAAYGFRKERALBQLQGHRIIVTLGVFTIHFSPNVPSCRLLLELDVSVSE 120
 DB 61 TTGAASAAYGFRKERALBQLQGHRIIVTLGVFTIHFSPNVPSCRLLLELDVSVSE 120
 QY 121 LLLYSHQCCSMMWIOHCARDVLEALAFIHHGYVHADLKPRNIIIMSAENECFKLIDFGL 180
 DB 121 LLLYSHQCCSMMWIOHCARDVLEALAFIHHGYVHADLKPRNIIIMSAENECFKLIDFGL 180
 QY 181 SPKEGNODVKYIQTODGYRAPEALONCLAQAGLSQDTECTSAVDLSLGIILLEMFSGMK 240
 DB 181 SPKEGN-DVKYIQTODGYRAPEALONCLAQAGLSQDTECTSAVDLSLGIILLEMFSGMK 239
 QY 241 LKHTVRSQEMKANSNAIIDIHIFASKAVVNAAIIPAYHLRDLIKSMHDDPSRRIIPAEALC 300
 DB 240 LKHTVRSQEMKANSNAIIDIHIFASKAVVNAAIIPAYHLRDLIKSMHDDPSRRIIPAEALC 299
 QY 301 SPFFSIPFAPHIEDLVMLPTPLRLNLVDDYLGNEBEYEDVVDVEECCQYGPVSVL 360
 DB 300 SPFFSIPFAPHIEDLVMLPTPLRLNLVDDYLGNEBEYEDVVDVEECCQYGPVSVL 359
 QY 361 LVPKENPGRGVFEVYANAGDSKAAOKLLTGRMFDGKPVVATFPLSAIKRGILYQTL 419
 DB 360 LVPKENPGRGVFEVYANAGDSKAAOKLLTGRMFDGKPVVATFPLSAIKRGILYQTL 418

RESULT 7
 ADD48093

ID ADD48093 standard; protein; 419 AA.

AC ADD48093;

DT 29-JAN-2004 (first entry)

DE Rat Protein AAC53031, SEQ ID NO 13791.

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urao D, Befort K, Costigan M;
 DR WPI; 2003-268312/26.

DR GENBANK; AAC53031.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX

PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 419 AA;

Query Match 99.1%; Score 2186; DB 7; Length 419;
Best Local Similarity 99.0%; Pred. No. 4,3e-228;
Matches 415; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGSSGGAMKAEPRPFEAFERGLMWQVSRIGSSGSSAVYRRCCGNGSPGALKQFLPFG 60
DB 1 MGSSGGAMKAEPRPFEAFERGLMWQVSRIGSSGSSAVYRRCCGNGSPGALKQFLPFG 60
QY 61 TTGAASAABEYGRKERRALEQOGHRNIVTLVGVFTIHPSPNVPSCLLLELDVSVSE 120
DB 61 TTGAASAABEYGRKERRALEQOGHRNIVTLVGVFTIHPSPNVPSCLLLELDVSVSE 120
QY 121 LLLYSSHQGCSMMWIOHCARDVLEALFLHHEGYVHADLPKRNILWSAENECFKLIDFGI 180
DB 121 LLLYSSHQGCSMMWIOHCARDVLEALFLHHEGYVHADLPKRNILWSAENECFKLIDFGI 180
QY 181 SPREGNOVDKYITODYRAPEALONCLAQAGQSTECTSANDLWSLIGILLMEFSGMK 240
DB 181 SPREGNOVDKYITODYRAPEALONCLAQAGQSTECTSANDLWSLIGILLMEFSGMK 240
QY 241 LKHTVASQEMKANSAAIIDIHIFASKAVVNAALIPAYHLRDLIKSLMLDDPSRRIPAEALC 300
DB 241 LKHTVASQEMKANSAAIIDIHIFASKAVVNAALIPAYHLRDLIKSLMLDDPSRRIPAEALC 300
QY 301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVEDVKEEQKYGFVSL 360
DB 301 SPFFSIPFAPHIEDLVMLPTPVRLINLVDDDYLGNEEYEDVEDVKEEQKYGFVSL 360
QY 361 LVPKEMPGQGVVEYANAGDSKAQKLTGMRPDKPVATFYPLSAKRGYLTTL 419
DB 361 LVPKEMPGQGVVEYANAGDSKAQKLTGMRPDKPVATFYPLSAKRGYLTTL 419

RESULT 8
ABP64662
ID ABP64662 standard; protein; 137 AA.

XX ABP64662;
AC
XX
XX 04-NOV-2002 (first entry)
DT
XX
XX Human ORF1032.

XX Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
XX Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
XX human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
XX cancer; cardiovascular disease; allergy; autoimmune disease;
XX wound healing; blood coagulation disorder; inflammatory disorder.

XX Homo sapiens.
OS
XX US2002082206-A1.
PN
XX 27-JUN-2002.
XX
XX
XX 30-MAY-2001; 2001US-00867550.
PF
XX 30-MAY-2000; 2000US-0208427P.
PR

XX (LEACH/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.

XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX WPI; 2002-626554/67.
DR N-PSDB; ABQ99225.
DR
XX

PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.

PS Claim 10; SEQ ID NO 2064; 78pp; English.

XX The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64661 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC segdata.uspto.gov/sequence.html?DocID=2002082206

XX Sequence 137 AA;

Query Match 21.9%; Score 483.5; DB 5; Length 137;
Best Local Similarity 80.0%; Pred. No. 1e-43;
Matches 96; Conservative 1; Mismatches 12; Indels 11; Gaps 1;

QY 63 GAAASAABEYGRKERRALEQOGHRNIVTLVGVFTIHPSPNVPSCLLLELDVSVSELL 122
DB 29 GGAAGAASGGGLAAVA-----VTLVGVFTIHPSPNVPSCLLLELDVSVSELL 77

QY 123 LYSSHQGCSMMWIOHCARDVLEALFLHHEGYVHADLPKRNILWSAENECFKLIDFGISF 182
DB 78 LYSSHQGCSMMWIOHCARDVLEALFLHHEGYVHADLPKRNILWSAENECFKLIDFGISF 137

RESULT 9
AAG38502
ID AAG38502 standard; protein; 638 AA.

XX AC AAG38502;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47508.
XX DE Protein identification; signal transduction pathway; metabolic pathway;
XX DE hybridization assay; genetic mapping; gene expression control; promoter;
XX DE termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-013149P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 16-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0148341P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.

```
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.0%; Score 243.5; DB 3; Length 638;
Best Local Similarity 25.9%; Pred. No. 1.4e-16;
Matches 106; Conservative 56; Mismatches 146; Indels 101; Gaps 17;
```

```
QY 11 EPPRFLEARGRLMVGQVSRLLGSSASVYRVRCGNGSPGALKQFLPPTGTAASAAAB 70
DB 249 DEDLFFKVKNGKLLQRLGKIGSGSSSEVHKYI--SSDCITTYALKKTKLKGKRDYATY---- 301
QY 71 YGFRKERAALEQIQGHRIIVTLVGVFTIHSPVVPSCRLLLELLDVSVB----- 120
DB 302 YGFCGQIGVILKXKLGKNTIIQL-----IDYEVTDKTLLQEVLTNGTMSNKGKRVKEDGF 354
QY 121 ---LLIYS---SHGGSNM-----MICHCAADVLEALAFLLHSGTYHADIK 160
DB 355 IYVAVLEYGEIDLAMHMSQKREIEGSDRTIDENWMLRFYMQQIIQAVNTIHEERIVHSDIK 414
QY 161 PNITLMSANECKKLLDFGLSPFEGNODVYIOTD-----GYRAPEAEIQNLAQAGIQ 214
DB 415 PNNFL--LYVRGFLKLLIDFGIA-KAINSDTTNIORDSQVGLTSTYSPFAWNCESDENG-- 469
```

```
QY 215 SDPECTSAVDLWLSGIIILEMFGMKLKHITVRSQEMKANSAAIID--HIFASKAVVNAAI 272
DB 470 NTIKGGRPSDIMSLSGCIILYQWVYG-RTPPADYKTFW-AKFXYITDPNHEITYNQLSNP-- 525
QY 273 PAVHRLDIKSMHLDDPSRRIRIEMALCSPPFSIFPAH----- 311
DB 526 ---WLIDMKKCLADNRQWRIRPEILOHPFLAPPI-PHEPQVKTIKLFSLIAESGSD 581
QY 312 -----IEDLVMLPPVRLRLNVLNVLDDDYLGNEEYEDVEDYKEEC 351
DB 582 DKANSMISQLEQLINPAPLPR-----NDVIDSDQNOQLSRVSELC 624

RESULT 10
AAG38501
ID AAG38501 standard; protein; 720 AA.
XX AAG38501;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47507.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137223P.
XX 03-JUN-1999; 99US-0137528P.
```

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142290P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159332P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.0%; Score 243.5; DB 3; Length 720;
Best Local Similarity 25.9%; Pred. No. 1.7e-16;

PR 22-JUL-1999; 99US-0145192P
PR 23-JUL-1999; 99US-0145145P
PR 23-JUL-1999; 99US-0145218P
PR 23-JUL-1999; 99US-0145224P
PR 26-JUL-1999; 99US-0145276P
PR 27-JUL-1999; 99US-0145913P
PR 27-JUL-1999; 99US-0145918P
PR 27-JUL-1999; 99US-0145919P
PR 28-JUL-1999; 99US-0145951P
PR 02-AUG-1999; 99US-0146386P
PR 02-AUG-1999; 99US-0146388P
PR 02-AUG-1999; 99US-0146389P
PR 03-AUG-1999; 99US-0147038P
PR 04-AUG-1999; 99US-0147204P
PR 04-AUG-1999; 99US-0147302P
PR 05-AUG-1999; 99US-0147192P
PR 05-AUG-1999; 99US-0147260P
PR 06-AUG-1999; 99US-0147303P
PR 06-AUG-1999; 99US-0147416P
PR 09-AUG-1999; 99US-0147493P
PR 09-AUG-1999; 99US-0147935P
PR 10-AUG-1999; 99US-0148171P
PR 11-AUG-1999; 99US-0148319P
PR 12-AUG-1999; 99US-0148341P
PR 13-AUG-1999; 99US-0148565P
PR 13-AUG-1999; 99US-0148584P
PR 16-AUG-1999; 99US-0149368P
PR 17-AUG-1999; 99US-0149175P
PR 18-AUG-1999; 99US-0149426P
PR 20-AUG-1999; 99US-0149722P
PR 20-AUG-1999; 99US-0149929P
PR 20-AUG-1999; 99US-0149929P
PR 23-AUG-1999; 99US-0149902P
PR 23-AUG-1999; 99US-0149930P
PR 25-AUG-1999; 99US-0150566P
PR 26-AUG-1999; 99US-0150884P
PR 27-AUG-1999; 99US-0151065P
PR 27-AUG-1999; 99US-0151066P
PR 27-AUG-1999; 99US-0151080P
PR 30-AUG-1999; 99US-0151303P
PR 31-AUG-1999; 99US-0151438P
PR 01-SEP-1999; 99US-0151930P
PR 07-SEP-1999; 99US-0152363P
PR 10-SEP-1999; 99US-0153070P
PR 13-SEP-1999; 99US-0153758P
PR 15-SEP-1999; 99US-0154018P
PR 16-SEP-1999; 99US-0154039P
PR 20-SEP-1999; 99US-0154779P
PR 22-SEP-1999; 99US-0155139P
PR 23-SEP-1999; 99US-0155486P
PR 24-SEP-1999; 99US-0155659P
PR 28-SEP-1999; 99US-0156458P
PR 29-SEP-1999; 99US-0156596P
PR 04-OCT-1999; 99US-0157117P
PR 05-OCT-1999; 99US-0157753P
PR 07-OCT-1999; 99US-0157855P
PR 07-OCT-1999; 99US-0158029P
PR 08-OCT-1999; 99US-0158232P
PR 12-OCT-1999; 99US-0158369P
PR 13-OCT-1999; 99US-0159293P
PR 13-OCT-1999; 99US-0159294P
PR 13-OCT-1999; 99US-0159295P
PR 14-OCT-1999; 99US-0159329P
PR 14-OCT-1999; 99US-0159330P
PR 14-OCT-1999; 99US-0159331P
PR 14-OCT-1999; 99US-0159637P
PR 14-OCT-1999; 99US-0159638P
PR 18-OCT-1999; 99US-0159584P
PR 18-OCT-1999; 99US-0160741P
PR 21-OCT-1999; 99US-0160767P
PR 21-OCT-1999; 99US-0160768P
PR 21-OCT-1999; 99US-0160770P
PR 21-OCT-1999; 99US-0160814P

PR 21-OCT-1999; 99US-0160815P
PR 22-OCT-1999; 99US-0160980P
PR 22-OCT-1999; 99US-0160981P
PR 22-OCT-1999; 99US-0160989P
PR 22-OCT-1999; 99US-0161040P
PR 25-OCT-1999; 99US-0161405P
PR 25-OCT-1999; 99US-0161405P
PR 25-OCT-1999; 99US-0161406P
PR 26-OCT-1999; 99US-0161359P
PR 26-OCT-1999; 99US-0161360P
PR 26-OCT-1999; 99US-0161361P
PR 28-OCT-1999; 99US-0161920P
PR 28-OCT-1999; 99US-0161922P
PR 28-OCT-1999; 99US-0161933P
PR 29-OCT-1999; 99US-0162142P

Query Match 11.0%; Score 243.5; DB 3; Length 777;
Best Local Similarity 25.9%; Pred. No. 1.9e-16;
Matches 106; Conservative 56; Mismatches 146; Indels 101; Gaps 17;

QY 11 EPPRFLEAFGRIMOVQSRIGSGSSASVYRRCRCNGPSPGALKQFLPPTGTGAASAAE 70
DB 388 DDPLEFFKVGKLYQRIGKIGSGSSSEVHKVI--SSDCTIVALKKIKLGRDYATA--- 440
QY 71 YGFRKERAALEQLQGRNIVTLVYGVFTIHFSBNVPSRCILLELDVSYSE----- 120
DB 441 YGFCOEIGYKTKLKKKNTNIQL-----IDYEVDKTLIQVLNGTSMNKGKRVKEDGF 493
QY 121 ---LLIYS---SHOCCSNW-----MIQHCARDVLEALAFLEHGYHADIK 160
DB 494 IYVIVLEYGEIDIALHMLSQKREIEGSDRTIDENMWRIFYWQITLQAVNTIHEERIYHSDIK 553
QY 161 PNNILMSANNEBCFKLIDPLSFKBGNQVYKIQTD-----GYRPEALQWCLAQAGQ 214
DB 554 PANFL--LVRGFLKLIIDFGLA-KAINSDTNIQRDSQVGTLSYMSPEAFMCNESPENG-- 608
QY 215 SDTECTSAVDLWLSGILLLEMFSGMKLKHITVASQCEMKANSSAIIID--HIFASKAVVNAI 272
DB 609 NTIKCGRPBIDWLSGLICILYQWYVG-RTPPADYKTFW-AKFKYITDPNHETITNOJLNP-- 664
QY 273 PAVHLRDLIKSMHDDPSRRIIPAEMALCSPFPSIPPAH----- 311
DB 665 ---WLIDLMKKCLAMDNRQWRIRIPELQHPFLAPP1-PHEPOVKTIKLFSLIAESGSD 720
QY 312 -----IEDLVMLPFPVLAFLAVLDDDYAGNEEVEDYEDYKEC 351
DB 721 DRANSMSIQLEQLLSNPAPLPR-----NDVLDSRDQNOQLRVSSELC 763

RESULT 12
AAM85028 standard; protein; 544 AA.
ID AAM85028;
XX AAM85028;
AC AAM85028;
XX 17-OCT-2003 (revised)
DT 08-FEB-1999 (first entry)
XX CDK2-green fluorescent protein usion product.
DE Human; CDK2 gene; fusion protein; green fluorescent protein; GFP;
XX intracellular signaling; chimera.
KM nequorea victoria.
OS Homo sapiens.
OS Chimeric.
XX MO9845704-A2.
XX 15-OCT-1998.
PD 07-APR-1998; 98WO-DK000145.
XX 07-APR-1998; 98WO-DK000145.
PF 07-APR-1997; 97DK-00000392.
XX

```
XX (NOVO ) NOVO-NORDISK AS.
PA Thastrup O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;
PI MPI; 1998-594491/50.
DR N-PSDB; AAV71073.
XX
XX Determining effect on signalling pathways in live cells from
PT redistribution of luminothores - specifically fusions of green
PT fluorescent protein with a signalling component, and new apparatus,
PT particularly for identifying toxins and potential therapeutic agents.
XX
XX Example 12; Page 208-209; 326pp; English.
XX
XX The present sequence represents a human CDK2-green fluorescent protein
XX (GFP) fusion protein. The fusion protein is used in an assay to exemplify
XX the invention. The specification describes how quantitative information
XX about the influence of a molecule on a cellular response is obtained by
XX recording the variation, caused by the molecule, on mechanically intact
XX living cells, in the spatially distributed light emitted from a
XX luminothore present in the cells. The variation in light emission is
XX processed to provide information that correlates spatial distribution to
XX the degree of the molecule. The method is used to identify agents that
XX (in)directly affect intracellular signalling, especially to screen for
XX potential therapeutic agents or toxins, and to identify new drug targets.
XX (Updated on 17-Oct-2003 to standardise OS field)
XX
XX Sequence 544 AA;
SQ
Query Match 10.9%; Score 239.5; DB 2; Length 544;
Best Local Similarity 27.0%; Pred. No. 3e-16;
Matches 112; Conservative 52; Mismatches 154; Indels 97; Gaps 18;
QY 23 MGVQSLRGSSSSASYRVRCGNGSPPGALKQF-LPFGTTGAASAAYGRKERAALR 81
DB 4 FQKVEKIGEGTGVVYKAR--NKLTEGVVALKKIRLDTEBGPSTAI-----REISLTK 56
QY 82 QLOGHNITVLYGVFTIHSFNVPSRCLLELDVSVSELLLYSSHOGSGMMIIOHCARD 141
DB 57 EL-NHNPNIYKLDV--IHTENKL---YLVEFLHQDKKMDASALTGIPPLIKSYLQ 110
QY 142 VLEALAFLLHHEGVVHADLKPRNIIWSAENECFKLIDFGLSFKEGNODVXY---IQTDGYR 198
DB 111 LLQGLAFCHSHRVLRDLKPNLLINTEG-AIKLADFGIARAFGVPRVTYTHEVVTLMWR 169
QY 199 APEAEIQLCLAQGLQSDTECTSAVDLSIGIILLEMFSGMKLKHTVRSQEKANSATI 258
DB 170 APEI-----LLGSKYSTAVIDWSLGCIFAEI-----VTRRALPGDSE--I 209
QY 259 DHIFASKAVVN-----AIPAYHL-----RDLIKSLMD 287
DB 210 DQLFRIFRLTGTDEVWVGVTSMFDYKSPFKMARQDPSKYVPLPDEGRSLLSQMLHY 269
QY 288 DPSRRIPAEMLCSPPFSIPFAPHIEDLVMLPTPVRLNLVLDLDYLGNEEYEDV 347
DB 270 DPKRISAKAALAHPPF-----QD-VTKVPVHLRLMD-----PVATWV 307
QY 348 KEKCCQKGYPVVSLVLPKENGPGQVEVEYANAGDSKAQKLTLGRMFDGKVVAT 402
DB 308 SKGEELFTGVVPLVBLDGVNGHKF-SVSGEGEGDGYGKLT-----LKFICTT 356
RESULT 13
ADD21398
ID ADD21398 standard; protein; 298 AA.
XX
XX ADD21398;
AC
XX
XX 15-JAN-2004 (first entry)
XX
XX Golden hamster cyclin dependent kinase (cdk) 2 protein.
XX
```

```
KM continual growth; cultured cell; cyclin dependent kinase; cdk4; cdk2;
KM cdk6; activating mutation; cell growth; cell division; cell cycle;
KM cancer-causing agent; continual growth-induced cell; enzyme;
KM golden hamster.
XX
XX Mesocricetus auratus.
XX
XX MO2003044169-A2.
XX
XX 30-MAY-2003.
XX
XX 15-NOV-2002; 2002MO-US036729.
XX
XX 15-NOV-2001; 2001US-0334760P.
XX
XX (UTEM ) UNIV TEMPE.
XX
XX Reddy PE, Rane SG, Mettus RV;
PI MPI; 2003-449813/42.
DR N-PSDB; ADD21399.
XX
XX A composition for reversibly inducing continual growth in normal cells
PT comprises a cyclin dependent kinase protein (e.g. cdk4, cdk2 or cdk6) or
PT its active fragment, derivative, homolog or analog, having an activating
XX mutation.
XX
XX Disclosure; Page 104; 77pp; English.
XX
XX This invention relates to a novel composition for inducing a reversible
XX state of a continual growth in cultured cells and comprises at least one
XX compound comprising a cyclin dependent kinase (cdk4, cdk2 or cdk6
XX protein having an activating mutation. Growth and division of living
XX cells involve a regular series of events and processes that comprises the
XX cell cycle. Cyclin dependent kinases cdk2, cdk4 and cdk6 are involved in
XX the control of G1, the point at which cells irrevocably commit to DNA
XX synthesis and thus enter the cell cycle. The invention is useful in
XX reversibly inducing continual growth in normal cells and may allow the
XX screening of cancer-causing agents with the continual growth-induced
XX cells. The present sequence is that of the golden hamster cdk2 protein
XX related to the invention. Note: Due to an error in the specification or
XX sequence listing, the Seq ID numbers given in the disclosure do not
XX correspond to those given in the sequence listing. It is therefore
XX unclear which Seq ID number corresponds to which sequence and exactly
XX which sequence is being claimed.
XX
XX Sequence 298 AA;
SQ
Query Match 10.8%; Score 238; DB 7; Length 298;
Best Local Similarity 29.3%; Pred. No. 1.7e-16;
Matches 99; Conservative 44; Mismatches 117; Indels 78; Gaps 16;
QY 23 MGVQSLRGSSSSASYRVRCGNGSPPGALKQF-LPFGTTGAASAAYGRKERAALR 81
DB 4 FQKVEKIGEGTGVVYKAR--NKLTEGVVALKKIRLDTEBGPSTAI-----REISLTK 56
QY 82 QLOGHNITVLYGVFTIHSFNVPSRCLLELDVSVSELLLYSSHOGSGMMIIOHCARD 141
DB 57 EL-NHNPNIYKLDV--IHTENKL---YLVEFLHQDKKMDASAVTGIPLPLIKSYLQ 110
QY 142 VLEALAFLLHHEGVVHADLKPRNIIWSAENECFKLIDFGLSFKEGNODVXY---IQTDGYR 198
DB 111 LLQGLAFCHSHRVLRDLKPNLLINTEG-IKLADFGIARAFGVPRVTYTHEVVTLMWR 169
QY 199 APEAEIQLCLAQGLQSDTECTSAVDLSIGIILLEMFSGMKLKHTVRSQEKANSATI 258
DB 170 APEI-LLGC-----KYSTAVIDWSLGCIFAEI-----VTRRALPGDSE--I 209
QY 259 DHIFASKAVVN-----AIPAYHL-----RDLIKSLMD 287
DB 210 DQLFRIFRLTGTDEVWVGVTSMFDYKSPFKMARQDPSKYVPLPDEGRSLLSQMLHY 269
QY 288 DPSRRIPAEMLCSPPFSIPFAPHIEDLVMLPTPVRL 325
```

```
Db 270 DENKRISAKALAHPEF-----OD-VTKKVPNRL 298
||::|||::|||
RESULT 14
AAE31916
ID AAE31916 standard; protein; 435 AA.
XX
XX AAE31916;
AC
XX 07-MAR-2003 (first entry)
XX
XX Yeast Hog1 mutant protein (Y68H).
DE
XX Yeast; Hog1; enzyme; mitogen activated protein kinase; MAPK pathway;
KM drug design; signal transduction; neoplasia; cancer; inflammation;
KM degenerative disease; immunological disorder; therapy; mutant; mtein.
XX
XX Saccharomyces cerevisiae.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 68 /note="Wild type Tyr substituted with His"
XX
XX MO200281746-A2.
XX
XX 17-OCT-2002.
XX
XX 04-APR-2002; 2002WO-IL000281.
XX
XX 05-APR-2001; 2001IL-00142481.
XX
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Engelberg D, Bell M, Levitzki A;
XX
XX MPI; 2003-067522/06.
XX
XX Screening for constitutively activated mutants of the mitogen activated
XX protein kinase (MAPK) pathway member, useful for screening drugs for
XX treating e.g. cancer, by employing a mutant yeast strain devoid of an
XX upstream kinase.
XX
XX Claim 22; Page; 123pp; English.
XX
XX The invention relates to a method of screening for constitutively active
XX mutant of a desired eukaryotic mitogen activated protein kinase (MAPK)
XX pathway member (such as Hog1 in yeast). The method involves employing a
XX mutant yeast strain devoid of an upstream kinase which is capable of
XX activating the MAPK pathway member of a yeast MAPK pathway that is
XX equivalent or corresponding to the desired eukaryotic MAPK pathway
XX member. The method is useful in drug design particularly for screening
XX constitutively active mutants of MAPKs and substances that are inhibitors
XX of MAPK signal transduction pathways. MAPK pathway inhibitors are used
XX for treating pathological conditions e.g. neoplasia, cancer,
XX inflammation, degenerative diseases or immunological disorders. The
XX present sequence is yeast Hog1 mutant protein. Note: This sequence is not
XX shown in the specification but is derived from the yeast Hog1 wild-type
XX protein shown as SEQ ID NO: 1 (AAE31909) in figure 3A of the
XX specification
XX
XX Sequence 435 AA;
XX
XX Query Match 10.7%; Score 236; DB 6; Length 435;
XX Best Local Similarity 25.7%; Pred. No. 5.1e-16;
XX Matches 98; Conservative 53; Mismatches 140; Indels 90; Gaps 18;
XX
XX 17 EAERGRIMQQR-----IGSGSSASVYRRCGMP--GSPPALQFLPFGTGAAGA 68
XX :||::|||::|||
XX 11 QIFGVFEITNRNDINPVGMGAFLV---CSATDTLSQPAIKIMKPFSTAVLAKR 66
XX :||::|||::|||
XX 69 AEYGFKEBAALBQLGHRNIVTLGVFTIHFPNVPSCRLLLELDVSVSELLVSSHQ 128
XX :||::|||::|||
```

```
Db 67 T-----HRELUKLKHLR-HEMLICLDIF---LSP-LEDIYFVTEHQGTDLHRLT---QTR 114
||::|||::|||
RESULT 15
AAB22137
ID AAB22137 standard; protein; 367 AA.
XX
XX AAB22137;
AC
XX 21-FEB-2001 (first entry)
XX
XX Rat stress-activated protein kinase 3.
XX
XX Rat; stress-activated protein kinase 3; SAPK3; PDZ domain; G1G2 loop;
XX DHF domain; muscular dystrophy; ERK6; p38gamma.
XX
XX Rattus sp.
XX
XX WO200048002-A1.
XX
XX 17-AUG-2000.
XX
XX 09-FEB-2000; 2000WO-GB000374.
XX
XX 09-FEB-1999; 99GB-00002696.
XX
XX (MED1-) MEDICAL RES COUNCIL.
XX
XX Spillantini MG, Goedert M, Hasegawa M, Buee-Scherrer V, Thomas G;
XX Cohen P, Cuenda A;
XX
XX MPI; 2000-558226/51.
XX
XX N-PSDB; AAA90869.
XX
XX Novel methods for identifying compounds useful for treating muscular
XX dystrophy comprising identifying modulators of protein kinase and PDZ
XX domain binding.
XX
XX Disclosure; Fig 6; 102pp; English.
XX
XX The present sequence is the protein sequence of the rat stress-activated
XX protein kinase 3 (SAPK3, also known as ERK6 and p38gamma) enzyme. It was
XX used to demonstrate the methods and peptides of the invention, which
XX involve the identification of compounds which are able to modulate the
XX interaction between PDZ domains (also known as DHF or G1G2 domains) and
XX protein kinases. The invention also comprises peptide fragments derived
XX from stress-activated protein kinases which are involved in protein
XX phosphorylation. The compounds and peptides can be used in the treatment
XX of muscular diseases, for example muscular dystrophy, or in cases where
XX modulation of phosphorylation of proteins with PDZ domains or modulation
XX of signalling via activin receptors or voltage gated channels is
XX necessary, or in instances where the localisation of SAPK3 at the
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:30 ; Search time 5310 Seconds

(without alignments)
8646.716 Million cell updates/sec

Title: US-10-798-532-3

Perfect score: 1260
Sequence: 1 atcgcgcgagctccgctgcgc.....tgatccaacctgtcttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|----------|--------------------|
| 1 | 1113.4 | 78.4 | 1763 | 3 | AK013347 | AK013347 Mus muscu |
| 2 | 886.6 | 70.4 | 2846 | 3 | AK030152 | AK030152 Mus muscu |
| 3 | 861.8 | 68.4 | 1079 | 4 | BM468107 | BM468107 AGENCOURT |
| 4 | 723.8 | 57.4 | 917 | 4 | BM451184 | BM451184 AGENCOURT |
| 5 | 622.8 | 49.4 | 1015 | 6 | BY712344 | BY712344 BY712344 |
| 6 | 619.2 | 49.1 | 657 | 2 | BG547675 | BG547675 602575520 |
| 7 | 617.8 | 49.0 | 889 | 2 | BE895119 | BE895119 601436041 |
| 8 | 616.4 | 48.9 | 772 | 4 | BI562420 | BI562420 603256316 |
| 9 | 608.8 | 48.3 | 932 | 5 | BU541306 | BU541306 AGENCOURT |
| 10 | 595 | 47.2 | 875 | 4 | BI259568 | BI259568 602968330 |
| 11 | 591.6 | 47.0 | 834 | 7 | CF593363 | CF593363 AGENCOURT |
| 12 | 583.8 | 46.3 | 906 | 4 | BG386905 | BG386905 602454704 |
| 13 | 583.6 | 46.3 | 682 | 4 | BI333581 | BI333581 602997069 |
| 14 | 578 | 45.9 | 888 | 5 | BQ225071 | BQ225071 AGENCOURT |
| 15 | 560 | 44.4 | 602 | 7 | CO737403 | CO737403 S1T02C17 |
| 16 | 557.6 | 44.3 | 779 | 5 | BU447831 | BU447831 603767571 |
| 17 | 556.4 | 44.2 | 876 | 5 | BU184346 | BU184346 AGENCOURT |
| 18 | 549.2 | 43.6 | 844 | 4 | BG474616 | BG474616 602517423 |
| 19 | 519.8 | 41.3 | 824 | 2 | BE568528 | BE568528 601347425 |
| 20 | 516.8 | 41.0 | 823 | 7 | CO771206 | CO771206 testis ES |
| 21 | 516 | 41.0 | 781 | 7 | CN233453 | CN233453 WB074D11 |
| 22 | 498.6 | 39.6 | 717 | 5 | BU385761 | BU385761 603582650 |
| 23 | 497.8 | 39.5 | 670 | 7 | CN359673 | CN359673 170004245 |
| 24 | 496.8 | 39.4 | 523 | 2 | BF929410 | BF929410 IL2-NT020 |

| | | | | | | |
|----|-------|------|------|---|----------|--------------------|
| 25 | 495.2 | 39.3 | 896 | 7 | CO773091 | CO773091 testis ES |
| 26 | 494.2 | 39.2 | 1003 | 5 | BQ922020 | BQ922020 AGENCOURT |
| 27 | 493 | 39.1 | 928 | 5 | BQ424438 | BQ424438 AGENCOURT |
| 28 | 482.4 | 38.3 | 716 | 7 | CF745774 | CF745774 UI-M-GVO- |
| 29 | 471.8 | 37.4 | 755 | 5 | BU71858 | BU71858 603568233 |
| 30 | 469.8 | 37.3 | 794 | 5 | BU205357 | BU205357 603104559 |
| 31 | 468.2 | 37.2 | 824 | 4 | BI253163 | BI253163 602973284 |
| 32 | 467.2 | 37.1 | 785 | 6 | CB990192 | CB990192 AGENCOURT |
| 33 | 463.6 | 36.8 | 870 | 5 | BQ216813 | BQ216813 AGENCOURT |
| 34 | 463.4 | 36.8 | 484 | 5 | BX281211 | BX281211 UI-M-FRO- |
| 35 | 463.4 | 36.8 | 699 | 5 | BU707284 | BU707284 UI-M-FRO- |
| 36 | 460 | 36.5 | 696 | 5 | BU439278 | BU439278 604147528 |
| 37 | 456.6 | 36.2 | 483 | 2 | BF929411 | BF929411 IL2-NT020 |
| 38 | 454.2 | 36.0 | 631 | 5 | BU112717 | BU112717 603131421 |
| 39 | 451.8 | 35.9 | 743 | 4 | BI562440 | BI562440 603256324 |
| 40 | 450.4 | 35.7 | 986 | 5 | BU751496 | BU751496 CH4#002.E |
| 41 | 449 | 35.6 | 933 | 2 | BE886040 | BE886040 601507152 |
| 42 | 433.2 | 34.4 | 820 | 7 | CK597557 | CK597557 AGENCOURT |
| 43 | 427.4 | 33.9 | 672 | 5 | BX261106 | BX261106 BX261106 |
| 44 | 420 | 33.3 | 604 | 5 | BX261107 | BX261107 BX261107 |
| 45 | 413.2 | 32.8 | 570 | 5 | BU293329 | BU293329 60416812 |

ALIGNMENTS

RESULT 1
AK013347
LOCUS
DEFINITION
AK013347 1763 bp mRNA linear HTC 03-APR-2004
Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:281045404 product:kinase interacting with
leukemia-associated gene (statmin), full insert sequence.

ACCESSION
AK013347
VERSION
AK013347.1 GI:12850651
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
1 Carninci, P. and Hayashizaki, Y.
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
2 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
3 Shiba, K., Itoh, M., Hayatsu, N., Sugahara, Y., Shiba, K.,
Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multiplexed sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research

|||||
Db 1239 CTGTGTCGAAGAAATCCTGGCAGAGGCAAGTCTTCGTGATGACGGAAAGCTGGG 1298
|||
Qy 1141 GATTCGAAGCTGGCGGAGAAATCTGACTGGAAGATGTTTATGGAAAGTTGTTG 1200
|||
Db 1299 GATTCGAAGCTGGCTGAGAAAGTTGACTGACGGAGAGATGTTTACGGAAAGTTGTTG 1358
|||
Qy 1201 GCTACATTCACCGCGTGAAGTCTCAAGAGAGGATATCTGTATCAAACTGCTTTAA 1260
|||
Db 1359 GCTACATTCACCGCGTGAAGTCTCAAGAGAGGATATCTGTATCAAACTGCTTTAA 1418
|||
RESULT 2 2846 bp mRNA linear HTC 03-Apr-2004
AK030152
LOCUS
DEFINITION
Mus musculus adult male testis cDNA, RIKEN full-length enriched
library, clone:4933401D07 product:kinase interacting with
leukemia-associated gene (statmin), full insert sequence.
AK030152
VERSION AK030152.1 GI:26326136
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
1 Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL 99279253
MEDLINE 10349636
PUBMED
REFERENCE
AUTHORS
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED
REFERENCE
AUTHORS
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Taahito, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujikake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL 20530913
MEDLINE 11076861
PUBMED
REFERENCE
AUTHORS
4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
6 (bases 1 to 2846)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, K., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
SUBMITTER Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
URL: http://location.qualifiers
FEATURES
source
1. 2846
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:4933401D07"
/db_xref="taxon:10090"
/clone="4933401D07"
/sex="male"
/tissue_type="testis"
/clone_id="RIKEN full-length enriched mouse cDNA library"
238..1230
/dev_stage="adult"
CDS
/note="unnamed protein product; kinase interacting with
leukemia-associated gene (statmin) (MGI:1341908)
putative"
/codon_start=1
/protein_id="BAC26812.1"
/db_xref="GI:26326137"
/translation="MTIVGVPIHSPVPSRCILILIDVSVSLILYSSHGCSMW
MIOHCRADYLEALFLHBYGTHADLKRNITMSKNECFKIDGLSFKEKNQVYK
IOTDYSRAVEALQWCLAQAGQSDTECSAVDLWSLITLLEMSGMLKRTVSQE
WANSASAIIDHIFASKAVVNAIPAYHDLKMLHDPGRIRIPAEMLCSPPSIP
FAPRHEIDLVLPPTPLRLINLVDDYLLNEDEDEVDVDECKQKGVVSLVPEK
NGRQGVFEVYNAGDSKAAQGLTGRMFDPGFVAATFPPLSAVRYLYOTLL"
2823..2828
/note="putative"
2846
/note="putative"
ORIGIN
polyA_signal
polyA_site
Query Match 70.4%; Score 886.6; DB 3; Length 2846;
Best Local Similarity 93.1%; Pred. No. 8.8e-247;
Matches 928; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
Qy 264 CATCGTACTTGTAAGAGTGTATTAATCAATCACTTTCTCCAAATGTCATCAGCTG 323
|||
Db 234 CTTCATGACTTTATATGAGAGTCTTATCACTTATCTCTCCATGTCATCAGCTG 293
|||
Qy 324 TCTGTTGCTTGAAGCTCTGATGTCAGTGTTCGGAATGCTCTTATATTCAGTCACCA 383
|||
Db 294 TCTGCTGCTTGAAGCTCTGATGTCAGTGTTCGGAATGCTCTTATATTCAGTCACCA 353
|||
Qy 384 GGGTGTTCATGATGATGATACACATTCGCGCCGAGATGTTTGGAGGCCCTGCTTT 443
|||
Db 354 GGGCTGCTCAGTGTGATATACAGACATGTCGCCGAGATGTTTGGAGGCCCTGCTTT 413
|||
Qy 444 TCTTCATCATGAGGCTATGTCATGCGGACTCAACACGTAACATATTTGAGAGTGC 503
|||
Db 414 TCTTCACATGAGGCTAGTCAATCAGACCTCAACACGAAACATCTCTGAGAGTGC 473
|||
Qy 504 AGAGAAATGATGTTTAAATCATTTGACTTTGACTTCAAAAGAGGAATCAGGA 563
|||
Db 474 GGAGAAATGATGTTTAAATCATTTGACTTTGACTTCAAAAGAGGAATCAGGA 533
|||

QY 1047 GGAGTGTCAAAATATGACCACTGTATCTTACTGTTCACAAAGAAATCTGGCAG 1106
 |||||
 DB 781 GGAGTGTCAAAATATGACCACTGTATCTTACTGTTCACAAAGAAATCTGGCAG 840
 |||||
 QY 1107 AGGACAAGTC-TTGTGTAGTATGCAAAATCTG-GTGATTCAAAGCTGGCAAAAT 1163
 |||||
 DB 841 AGGACAAGTC-TTGTGTAGTATGCAAAATCTGNGTGATTCAAAGCTGGCAAAAT 899
 |||||
 RESULT 4
 BM451184 917 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT 6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
 DEFINITION 5', mRNA sequence.
 ACCESSION BM451184
 VERSION BM451184.1 GI:18500224
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 917)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsaps-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: LIML12121 row: j column: 24
 High quality sequence stop: 718.
 Location/Qualifiers
 1. 917
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5495159"
 /issue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: PCW-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."
 ORIGIN
 Query Match 57.4%; Score 723.8; DB 4; Length 917;
 Best Local Similarity 99.7%; Pred. No. 2e-199;
 Matches 725; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 261 TGATCATATTTGCCAGTAAGACAGTGTGATGCGCAATTCAGCTATCACCTAAG 320
 |||||
 QY 834 AACCTTATCAAAAGATCTTCATGATATGATCCAGAGACAAAGATTCCTGGAATGCG 893
 |||||
 DB 321 AGACCTTATCAAAAGATCTTCATGATATGATCCAGAGACAAAGATTCCTGGAATGCG 380
 |||||
 QY 894 ATTGTCAGCCCATCTTATGATTCCTTTGCCCTCATATTTGAAGATCTGGTCAATGCT 953
 |||||
 DB 381 ATTGTCAGCCCATCTTATGATTCCTTTGCCCTCATATTTGAAGATCTGGTCAATGCT 440
 |||||
 QY 954 TCCACTCCAGTCTTAAGATCTGCTGATGATGCTGATGATTTATCTTGGAAATGAGA 1013
 |||||
 DB 441 TCCACTCCAGTCTTAAGATCTGCTGATGATGCTGATGATTTATCTTGGAAATGAGA 500
 |||||
 QY 1014 GGAATATGAAGATCTTGTGAAGATGTAAAGAGAGTGTCAAAATATGACCAATGCT 1073
 |||||
 DB 501 GGAATATGAAGATCTTGTGAAGATGTAAAGAGAGTGTCAAAATATGACCAATGCT 560
 |||||
 QY 1074 ATCTCTACTTGTTCACAAAGAAATCTGGCAGAGACAAAGTCTTGTGAGTATGACAA 1133
 |||||
 DB 561 ATCTCTACTTGTTCACAAAGAAATCTGGCAGAGACAAAGTCTTGTGAGTATGACAA 620
 |||||
 QY 1134 TGTGTGTATTCACAAAGCTGCGCAGAAATTAAGTCTGAGAGATGTTGATGGAGTT 1193
 |||||
 DB 621 TGTGTGTATTCACAAAGCTGCGCAGAAATTAAGTCTGAGAGATGTTGATGGAGTT 680
 |||||
 QY 1194 TGTGTGTATTCACAAAGCTGCGCAGAAATTAAGTCTGAGAGATGTTGATGGAGTT 1253
 |||||
 DB 681 TGTGTGTATTCACAAAGCTGCGCAGAAATTAAGTCTGAGAGATGTTGATGGAGTT 740
 |||||
 QY 1254 GCTTTAA 1260
 |||||
 DB 741 GCTTTAA 747
 |||||
 RESULT 5
 BY712344 1015 bp mRNA linear EST 17-DEC-2002
 LOCUS RIKEN full-length enriched, 10, 11 days embryo whole body
 DEFINITION Mus musculus cDNA clone 2810454004 5', mRNA sequence.
 ACCESSION BY712344
 VERSION BY712344.1 GI:27123621
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 1015)
 AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gotohori, T., Baldairelli, R., Hill, D.P., Bulic, C.,
 Hume, D.A., Quackenbush, J., Schrim, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,
 Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, J., Jarvis, B.L., Kanaya, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Malais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.U., Pereira, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shinada, K.,
 Sultana, R., Takemura, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itch, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

JOURNAL MEDLINE PUBLISHED

12466851
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saitama-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itch, M., Kagawa, I., Kawai, J., Kojima, Y., Kondou, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohnishi, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES

source
 Location/Qualifiers
 1. 1015
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="2810454004"
 /issue_type="whole body"
 /dev_stage="10, 11 days embryo"
 /clone_lib="RIKEN full-length enriched, 10, 11 days embryo whole body"

ORIGIN

Query Match 49.4%; Score 622.8; DB 6; Length 1015;
 Best Local Similarity 90.0%; Pred. No. 6.3e-170;
 Matches 678; Conservative 0; Mismatches 73; Indels 2; Gaps 1;
 1 ATGGCGGAGATCCGGCTGGCGTGGCGGCGGAGACCGCGGCTTTCTTGAGAGCGCTTCGGG 60
 159 ATGGCGGAGATCCGGCTGGCGTGGCGGCGGAGACCGCGGCTTTCTTGAGAGCGCTTCGGG 218
 61 CGGCTGTGGAGATACAGAGCGCTGTGGTGGCGGCTTCGCGCTCGGATGTGCGGGTT 120
 219 CGGCTGTGGAGATACAGAGCGCTGTGGTGGCGGCTTCGCGCTCGGATGTGCGGGTT 278
 121 CGGCTGTGGAGATACAGAGCGCTGTGGTGGCGGCTTCGCGCTCGGATGTGCGGGTT 180
 279 CGGCTGTGGAGATACAGAGCGCTGTGGTGGCGGCTTCGCGCTCGGATGTGCGGGTT 338
 181 ACCACCGGGGCTGGCGCTTCGCGGCGGATGTGTTTCCGCAAGAGAGGGGGCGCTG 240

339 ACCACCGGGGCTGGCGCTTCGCGGCGGAGATGTGTTTCCGCAAGAGAGGGCGGCTG 398
 241 GAACAGTTGACAGGCTCACAGAAACATCGTACTTTGTATGAGTGTTTACATTCACCTTT 300
 399 GANCAAGTTGACAGGCTCACAGAAACATCGTACTTTGTATGAGTGTTTACATTCACCTTT 458
 301 TCTCCAAATGTGCAATCAGGCTGTCTGTGCTTGAATCTCTGGATGTCAAGTCTTTCGGA 360
 459 TCTCCAAATGTGCAATCAGGCTGTCTGTGCTTGAATCTCTGGATGTCAAGTCTTTCGGA 518
 361 TTGCTCTTATTTTCAAGTACAGGAGGTTTTCATGTGGATGTACAGATTTGCGCCGA 420
 519 TTGCTCTTATTTTCAAGTACAGGAGGTTTTCATGTGGATGTACAGATTTGCGCCGA 578
 421 GATGTTTGGAGCGCTTCTTCTTCTTATCATCATGAGGCTATGTCCATGCGGACCTCAA 480
 579 GATGTTTGGAGCGCTTCTTCTTCTTATCATCATGAGGCTATGTCCATGCGGACCTCAA 638
 481 CCAAGTACATATTTGTGAGTGCAGAGAAATGATTTTAAATCATTTGACTTTGACTT 540
 639 CCACGAAACATCTGTGTGAGTGCAGAGAAATGATTTTAAATCATTTGACTTTGACTT 698
 541 AGCTTCAAGAGGCAATCAGAGTGTAAAGTATTTATTCAGACAGAGGGATGCGGCTCCA 600
 699 AGCTTCAAGAGGCAATCAGAGTGTAAAGTATTTATTCAGACAGAGGGATGCGGCTCTT 758
 601 GAAGCAGATTTGCAAAATTTGCTTGGCCAGAGCTGCTGCAGAGTATGATGATGATGAC 660
 759 GAAGCAGAGCTGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
 661 TGAAGTGTGATGTGTGAGAGCTGAGAGTATCATTTTACTGGAATTTCTCAGGAATGAA 720
 819 TGAAGTGTGATGTGTGAGAGCTGAGAGTATCATTTTACTGGAATTTCTCAGGAATGAA 878
 721 CTGAAACATACAGTACAGATCTCAGGAATGGAAG 753
 879 CTGAA--AATACATACAGTCTCAGGAATGGAAG 909

RESULT 6
 BG547675 657 bp mRNA linear EST 04-APR-2001
 LOCUS 602575520P1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703645 5',
 DEFINITION mRNA sequence.
 ACCESSION BG547675
 VERSION BG547675.1 GI:13546340
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@bbs-rm1.nhl.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at:
 http://image.llnl.gov
 Plate: L10M1542 row: g column: 06
 High quality sequence stop: 657.
 Location/Qualifiers
 1. 657
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4703645"
 /lab_host="DH10B (TI phage-resistant)"

/clone.lib="NIH_MGC_77"
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
 SfiI (ggcgccctggcc); Site 2: SfiI (ggcgatcatggc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGCCGAGCGCCGAGC-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.9
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 49.1%; Score 619.2; DB 4; Length 657;
 Best Local Similarity 99.1%; Pred. No. 6.2e-169;
 Matches 654; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 577 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGCTGGCCAGGCTGGC 636
 DB 1 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGCTGGCCAGGCTGGC 60
 QY 637 CTGCAAGAGTATCAGAAATGTACTCACTGTTGATCTGTGGAGCCTAGGAATCATTTTA 696
 DB 61 CTGCAAGAGTATCAGAAATGTACTCACTGTTGATCTGTGGAGCCTAGGAATCATTTTA 120
 QY 697 CTGGAATATGTTCTCAGAAATGAAACTGAAACATACAGTCAATCTAGGAATGGAAGGCA 756
 DB 121 CTGGAATATGTTCTCAGAAATGAAACTGAAACATACAGTCAATCTAGGAATGGAAGGCA 180
 QY 757 AACAGTTCTGCTATTATTGATCAATATTTGACAGTAAGCAGTGGTGAATGCCGCAATT 816
 DB 181 AACAGTTCTGCTATTATTGATCAATATTTGACAGTAAGCAGTGGTGAATGCCGCAATT 240
 QY 817 CCAAGCTTACCTTAAGAGACTTATCAAAAGCATGCTTCATGATGATCCAGACAGAGA 876
 DB 241 CCAAGCTTACCTTAAGAGACTTATCAAAAGCATGCTTCATGATGATCCAGACAGAGA 300
 QY 877 ATTCCGCTGAATGAGCATTTGTCAGCCCAATCTTGAATCTCTTGGCCCTCATATT 936
 DB 301 ATTCCGCTGAATGAGCATTTGTCAGCCCAATCTTGAATCTCTTGGCCCTCATATT 360
 QY 937 GAAGATCTGATCAGTCTCCACTCCAGTCTAAGACTGCTGAATGCTGATGATGAT 996
 DB 361 GAAGATCTGATCAGTCTCCACTCCAGTCTAAGACTGCTGAATGCTGATGATGAT 420
 QY 997 TATCTTGGAAATGAAGAAATATGAATGTTGTAAGAGATGTAAGAAGAGTCAA 1056
 DB 421 TATCTTGGAAATGAAGAAATATGAATGTTGTAAGAGATGTAAGAAGAGTCAA 480
 QY 1057 AATATNGACCAAGTATCTCTACTGTTCCAAAGAAATCTCTGCGACAGACAAATC 1116
 DB 481 AATATNGACCAAGTATCTCTACTGTTCCAAAGAAATCTCTGCGACAGACAAATC 540
 QY 1117 TTGTGATGATGCAAAATCTGCTGATTCCAAAGCTGCGACAGAAATTAATCTGAAGAAG 1176
 DB 541 TTGTGATGATGCAAAATCTGCTGATTCCAAAGCTGCGACAGAAATTAATCTGAAGAAG 599
 QY 1177 ATGTTGATGGAAGTTGTTGCTGCTACATTTACCCGCTGAGTGCCTTACAGAGGGA 1236
 DB 600 ATGTTGATGGAAGTTGTTGCTGCTACATTTACCCGCTGAGTGCCTTACAGAGGGA 657

RESULT 7
 BE895119 899 bp mRNA linear EST 20-OCT-2000
 LOCUS 60136041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5',
 DEFINITION mRNA sequence.
 ACCESSION BE895119
 VERSION BE895119.1 GI:10358193
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 1 (bases 1 to 899)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LMA9753 row: b column: 12
 High quality sequence stop: 626.
 Location/Qualifiers

FEATURES
 source
 1..899
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3920939"
 /issue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone.lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 49.0%; Score 617.8; DB 2; Length 899;
 Best Local Similarity 99.7%; Pred. No. 1.8e-168;
 Matches 619; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 337 CTCTGAGTGTGAGTTCGGAATTCCTTAATTCAGTCAACAGGTTGTCATG 396
 DB 1 CTCTGAGTGTGAGTTCGGAATTCCTTAATTCAGTCAACAGGTTGTCATG 60
 QY 397 TGGATGATACAGCATTCGCGCCGAGATGTTTGAAGCCCTTCTTTCATCATGAG 456
 DB 61 TGGATGATACAGCATTCGCGCCGAGATGTTTGAAGCCCTTCTTTCATCATGAG 120
 QY 457 GCGTATGTCATGCGGACCTCAACACGTAACTATTTGAGTGCAGAAATGAAT 516
 DB 121 GCGTATGTCATGCGGACCTCAACACGTAACTATTTGAGTGCAGAAATGAAT 180
 QY 517 TTTAACTCATTTGACTTTGACTTGAAGCTTCAAGAGCAATCAGATGTAAGTATAT 576
 DB 181 TTTAACTCATTTGACTTTGACTTGAAGCTTCAAGAGCAATCAGATGTAAGTATAT 240
 QY 577 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGCTGGCCAGGCTGGC 636
 DB 181 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGCTGGCCAGGCTGGC 300
 QY 241 CAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGCTGGCCAGGCTGGC 300
 DB 637 CTGCAAGAGTATCAGAAATGTACTCACTGTTGATCTGTGGAGCCTAGGAATCATTTTA 696
 QY 301 CTGCAAGAGTATCAGAAATGTACTCACTGTTGATCTGTGGAGCCTAGGAATCATTTTA 360
 DB 697 CTGGAATATGTTCTCAGAAATGAAACTGAAACATACAGTCAATCTAGGAATGGAAGGCA 756
 DB 361 CTGGAATATGTTCTCAGAAATGAAACTGAAACATACAGTCAATCTAGGAATGGAAGGCA 420
 QY 757 AACAGTTCTGCTATTATTGATCAATATTTGACAGTAAGCAGTGGTGAATGCCGCAATT 816
 DB 421 AACAGTTCTGCTATTATTGATCAATATTTGACAGTAAGCAGTGGTGAATGCCGCAATT 480
 QY 817 CCAAGCTTACCTTAAGAGACTTATCAAAAGCATGCTTCATGATGATCCAGACAGAGA 876
 DB 481 CCAAGCTTACCTTAAGAGACTTATCAAAAGCATGCTTCATGATGATCCAGACAGAGA 540
 QY 877 ATTCCGCTGAATGAGCATTTGTCAGCCCAATCTTGAATCTCTTGGCCCTCATATT 936

```

Db      541 ATTCCTGTGAATAGGATGTGACAGCCATCTTTCAGATTCCTTTGCCCCCATATT 600
Qy      937 GAAGATCTGATGATGCTCC 957
Db      601 GAAGATCTGATGATGCTTCC 621

RESULT 8
B1562420      772 bp      mRNA      linear      EST 05-SEP-2001
LOCUS      603256316F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
DEFINITION      mRNA sequence.
ACCESSION      B1562420
VERSION      B1562420
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE      NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-r@mail.nih.gov
              Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
              CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shuraki
              Toshiyuki and Piero Carninci (RIKEN)
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              plate: LLAM11755 row: p column: 03
              High quality sequence stop: 727.
              Location/Qualifiers
                1. 772
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5298674"
                  /lab_host="DH10B"
                  /clone_lib="NIH_MGC_97"
                  /note="Organ: testis; Vector: pBluescript (modified
                  pBluescript KS+; Site 1: BamHI; Site 2: SalI-XhoI
                  (gagcag); Oligo-dT primed using primer
                  5'-TTTTTTTTTTTTTTVN-3', size selected for average
                  insert size 2.2 kb and normalized to 10^5. This is a
                  primary library enriched for full-length clones and
                  constructed using the Cap-trapper method (Carninci, in
                  preparation). Library constructed by M. Brownstein
                  (NHGRI/NHGRI, National Institutes of Health). Note: this is
                  a NIH_MGC Library."

ORIGIN

Query Match      48.9%; Score 616.4; DB 4; Length 772;
Best Local Similarity 99.4%; Pred. No. 4.3e-168;
Matches 628; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy      274 TTGTATGAGATGTTTACATTCACATTTCTTCACAAATGTCACAGGCTGTCTTCTT 333
Db      108 TTGTATGAGATGTTTACATTCACATTTCTTCACAAATGTCACAGGCTGTCTTCTT 167
Qy      334 GAACCTCGATGTCAGTGTGGGAATGCTCTTATATTCAGTACACAGGATGTTC 393
Db      168 GAACCTCGATGTCAGTGTGGGAATGCTCTTATATTCAGTACACAGGATGTTC 227
Qy      394 ATGTGATGATAGACATTCGCCGCCAGATGTTTGGAGGCCCTTGCTTTCTTCATCAT 453
Db      228 ATGTGATGATAGACATTCGCCGCCAGATGTTTGGAGGCCCTTGCTTTCTTCATCAT 287
Qy      454 GAGGGCTATGTCATGCGGACCTCAACAGTAACTATATTTGGAGTGCAGAGATGAA 513

```

```

Db      288 GAGGCTATGTCATCGGAGCTCAACCACTAATATTGTGAGTGCAGAGATGAA 347
Qy      514 TGTTTTAAATCATGATGATTTGATGATGATTTCAAGAGGCAATCAGATGTAAGTAT 573
Db      348 TGTTTTAAATCATGATGATTTGATGATGATTTCAAGAGGCAATCAGATGTAAGTAT 407
Qy      574 ATTCAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGATGCTTGCCCAAGCT 633
Db      408 ATTCAGACAGACGGGTATCGGGCTCCAGAGCAGAAATTCGATGCTTGCCCAAGCT 467
Qy      634 GGCCTGCAGAGATGATACAGAAATGTAACCTCAGCTGTGATCTGTGAGGCTAGATCAT 693
Db      468 GGCCTGCAGAGATGATACAGAAATGTAACCTCAGCTGTGATCTGTGAGGCTAGATCAT 527
Qy      694 TTAGTGAATATGTTCTCAGATGATGAAATGAAACATAGATCAGATCAGAGATGAAG 753
Db      528 TTAGTGAATATGTTCTCAGATGATGAAATGAAACATAGATCAGATCAGAGATGAAG 587
Qy      754 GCAACAGTTCGTCTATTATTTGATACA-TATTTGCCAGTAAAGCAGTGTGAATGCCG 812
Db      588 GCAACAGTTCGTCTATTATTTGATACA-TATTTGCCAGTAAAGCAGTGTGAATGCCG 647
Qy      813 AATTCAGCTATACCTTACCTAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAG 872
Db      648 AATTCAGCTATACCTTACCTAGAGACCTTATCAAAAGCATGCTTCATGATGATCCAAGCAG 707
Qy      873 AAGAAATCCTGCTGAATGTCATGTGACGCC 904
Db      708 AAGAAATCCTGCTGAATGTCATGTGACGCC 729

```

```

RESULT 9
B0541306      932 bp      mRNA      linear      EST 13-SEP-2002
LOCUS      AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
DEFINITION      5', mRNA sequence.
ACCESSION      B0541306
VERSION      B0541306.1 GI:22851747
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE      NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs-r@mail.nih.gov
              Tissue Procurement: DCTD/DRP
              CDNA Library Preparation: Rubin Laboratory
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              plate: L1CM2765 row: o column: 18
              High quality sequence stop: 631.
              Location/Qualifiers
                1. 932
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6572970"
                  /rname="type="carcinoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_40"
                  /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
                  Site 2: EcoRI; cDNA made by oligo-dT priming.
                  Directionally cloned into EcoRI/XhoI sites using the
                  following 5' adaptor: GGCACGAG(G). Library constructed by
                  Ling Hong in the laboratory of Gerald M. Rubin (University
                  of California, Berkeley) using ZAP-cDNA synthesis kit

```

FEATURES

source

QY 897 GTGACAGCCATCTTTAGCATTCCTTTTGGCCCCCATATTGAAGATCT-GGTGATGCTTC 955
 |||||
 DB 593 GTGACAGCCATCTTTAGCATTCCTTTTGGCCCCCATATTGAAGATCTGGGTATGCTTC 652
 |||||
 QY 956 CCACTCCAGTG--CTAAGACTGCTGAATGCTGTGATGATGATTAATCTGGAGATGA--G 1012
 |||||
 DB 653 CCACCTCAGAGGGCTAAAGACTGCTGAATGCTGTGATGATGATTAATCTGGAGATGAAG 712
 |||||
 QY 1013 AGGAATATGAG--ATGTTGAGAG--ATGTAAAGAGAGGTGCAAAAATATGACACAG 1069
 |||||
 DB 713 AGGAATATGAGATGTTGAGAGATGTAAGAGAGAGGTGCAACCAATTTGACCA 772
 |||||
 QY 1070 TGGTATCTACTTGT-----CCAAAGGAAATCTGGCAGAGACA-AGTCTTTGTT 1122
 |||||
 DB 773 GGTGATATCTACTTGTTCCTCAAAAAGGAAATCTGGGCGAGAGACAGCTTTGTA 832
 |||||
 QY 1123 G-AGTATGCAAT-GCTGTGATTCCAAGCTGCGCAAAATT 1163
 |||||
 DB 833 GAAATATGCAATGTGGGGATTCCAAAAGTGGCGAAATT 875
 |||||

RESULT 11 834 bp mRNA linear EST 26-SEP-2003
 CFS93363
 LOCUS AGENCOURT 15622279 NIH_MGC_147 Homo sapiens cDNA clone
 DEFINITION IMAGE:30531268 5', mRNA sequence.
 CFS93363
 ACCESSION CFS93363.1 GI:36346826
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 QY 1 (bases 1 to 834)
 NIH-MGC http://mgc.nci.nih.gov/
 DB National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Daniela S. Gerhard, Ph.D.
 ORGANISM National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgsdbs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM621 row: m column: 05
 High quality sequence stop: 660.
 Location/Qualifiers

FEATURES
 source

1. 834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30531268"
 /issue_type="Human Placenta"
 /lab_host="DH10B Tona"
 /clone_1b="NIH_MGC_147"
 /note="Organ: placenta; Vector: pBluescript; Site 1:
 alt-XhoI; Site 2: BamH; Oligo-dT primed using primer
 5'-TTTTTTTTTTTNN-3', size-selected for average
 insert size 2.3 kb and normalized to 10⁵. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: This is
 a NIH_MGC library."

ORIGIN

Query Match 47.0%; Score 591.6; DB 7; Length 834;
 Best Local Similarity 97.4%; Pred. No. 7,6e-161;
 Matches 644; Conservative 0; Mismatches 10; Indels 7; Gaps 4;

QY 1 ATGGGGGATTCGGGCTGGGGCTGGGGCGGAGCGCGCGTTTCTGAGGCTTCGGG 60
 |||||
 DB 155 ATGGGGGATTCGGGCTGGGGCTGGGGCGGAGCGCGCGTTTCTGAGGCTTCGGG 214
 |||||
 QY 61 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCGCTCGGTGTATCGGGTT 120
 |||||
 DB 215 CGGCTGTGGCAGGTACAGAGCCGTCTGGGTAGCGGCTCCGCTCGGTGTATCGGGTT 274
 |||||
 QY 121 CGCTGCTGGCAGACCTTGGCTGCGCCCGGCGCTTCAGCACTTCTTCCGCGAGGA 180
 |||||
 DB 275 CGCTGCTGGCAGACCTTGGCTGCGCCCGGCGCTTCAGCACTTCTTCCGCGAGGA 334
 |||||
 QY 181 ACCACCGGGGCTGGCGCTTCGCGCGAGTATGTTTCCGAAAGAGAGGGCGGCTG 240
 |||||
 DB 335 ACCACCGGGGCTGGCGCTTCGCGCGAGTATGTTTCCGAAAGAGAGGGCGGCTG 394
 |||||
 QY 241 GAACAGTTGACAGGTACAGAAACATGTAATTTGTATGAGTGTATCAATCCACTTT 300
 |||||
 DB 395 GAACAGTTGACAGGTACAGAAACATGTAATTTGTATGAGTGTATCAATCCACTTT 454
 |||||
 QY 301 TCTCAATATGTCATCAAGCTGTCTGTTGCTTGAATCTCTGAGATGTCAGTGTTCGGA 360
 |||||
 DB 455 TCTCAATATGTCATCAAGCTGTCTGTTGCTTGAATCTCTGAGATGTCAGTGTTCGGA 514
 |||||
 QY 361 TTGCTCTTATTTCCAGTACACAGGGTTGTTCCATGTGGATGATACAGATTGGCGGGA 420
 |||||
 DB 515 TTGCTCTTATTTCCAGTACACAGGGTTGTTCCATGTGGATGATACAGATTGTGCGGA 574
 |||||
 QY 421 GATGTTTGGAGCCCTTCTTCTTCATCAATGAGGGCTATGTCATGCGGACTCAAA 480
 |||||
 DB 575 GATGTTTGGAGCCCTTCTTCTTCATCAATGAGGGCTATGTCATGCGGACTCAAA 634
 |||||
 QY 481 CCACTGATACATTTTGTGAGTGCAGAGATGAATGTTTAACTATGATCTTGGACTT 540
 |||||
 DB 635 CCACTGATACATTTTGTGAGTGCAGAGATGAATGTTTAACTATGATCTTGGACTT 694
 |||||
 QY 541 A-GCTTCAAGAGAGCAATCAGATGTAAAGTATTCAGACAGCGGATATC-GGGCTC 598
 |||||
 DB 695 AGGCTTCAGAGAGCAATCAGATGTAAAGTATTCAGACAGCGGATATC-GGGCTC 754
 |||||
 QY 599 CAGAAGCAGAAATTCAGAAATTCCTT---GGCCAGGCT--GGCTGCAGAGTATCAGA 653
 |||||
 DB 755 CAGAAGCAGAAATTCAGAAATTCCTTGGGCCCCAGGCTTGGCTGCAGAGTATCAGA 814
 |||||
 QY 654 A 654
 |||||
 DB 815 A 815
 |||||

RESULT 12 906 bp mRNA linear EST 12-MAR-2001
 BG386905
 LOCUS 602454704P1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4583065 5',
 DEFINITION mRNA sequence.
 ACCESSION BG386905 GI:13280454
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 QY 1 (bases 1 to 906)
 NIH-MGC http://mgc.nci.nih.gov/
 DB National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgsdbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: NIH Intramural Sequencing Center
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLCMI306 row: 9 column: 02
 High quality sequence stop: 668.

FEATURES

source

Location/Qualifiers

1..906
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4583065"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_15"
 /note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACCGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)"

ORIGIN

Query Match

46.3%; Score 583.8; DB 4; Length 906;

Best Local Similarity 96.1%; Pred. No. 1.5e-158;

Matches 664; Conservative 0; Mismatches 17; Indels 10; Gaps 6;

QY 570 GATATTTCAGACGAGGGGATCGGGCTCAGAGCAAGCAATTCATTCGGCCCA 629
 Db 2 GTATATTCAGACGAGGGGATCGGGCTCAGAGCAAGCAATTCATTCGGCCCA 61
 QY 630 GGCTGGCTGACAGATGATACAGATGATCCTCAGCTGTTGATCTGAGCCTTAGAAT 689
 Db 62 GGCTGGCTGACAGATGATACAGATGATCCTCAGCTGTTGATCTGAGCCTTAGAAT 121
 QY 690 CATTTTACGGAATGTTCTCAGAGTAAGAACTGAAACATACGTCAGATCTCAGAAATG 749
 Db 122 CATTTTACGGAATGTTCTCAGAGTAAGAACTGAAACATACGTCAGATCTCAGAAATG 181
 QY 750 GAAGCAACAGTTCGCTATTTATGATCATATTTGGCAGTAAGACAGTGTGAATGC 809
 Db 182 GAAGCAACAGTTCGCTATTTATGATCATATTTGGCAGTAAGACAGTGTGAATGC 241
 QY 810 CGCAATTCAGCCTTACCTTAAGAGACCTTATCAAAAGCATCTTCATGATGATCAAG 869
 Db 242 CGCAATTCAGCCTTACCTTAAGAGACCTTATCAAAAGCATCTTCATGATGATCAAG 301
 QY 870 CAGAGAAATTCCTGCTGAAATGCGATTGCGACCCCATTTTGAATTCCTTTGCCCC 929
 Db 302 CAGAGAAATTCCTGCTGAAATGCGATTGCGACCCCATTTTGAATTCCTTTGCCCC 361
 QY 930 TCATATTTGAAGATCTGTCATGCTTCCACTCCAGTGTGAAGATCTGTAATGTGCGGA 989
 Db 362 TCATATTTGAAGATCTGTCATGCTTCCACTCCAGTGTGAAGATCTGTAATGTGCGGA 421
 QY 990 TGAATGATTTCTTGGAAATGAAGAGAAATGAAGATGTTGAAGATGTAAAGAGGA 1049
 Db 422 TGAATGATTTCTTGGAAATGAAGAGAAATGAAGATGTTGAAGATGTAAAGAGGA 481
 QY 1050 GTGTCAAAAATATGACCAAGTGTATCTCTAATTTTCCAAAGAAAATCTTGCGAGAG 1109
 Db 482 GTGTCAAAAATATGACCAAGTGTATCTCTAATTTTCCAAAGAAAATCTTGCGAGAG 539
 QY 1110 ACAAGCTTTGTTGAGATCAATGCTGCTGATTCCTCAAAAGTGGCGAGAAATTAACGAC 1169
 Db 540 ACAAGCTTTGTTGAGATCAATGCTGCTGATTCCTCAAAAGTGGCGAGAAATTAACGAC 598
 QY 1170 TGAAGAGATTTGATGGAAGTTTGTGTGCTACATTCACCGCTGAGTGCCTACAA 1229
 Db 599 TGAAG-ATGTTGATGGAAGTTTGTGTGCTACATTCACCGCTGAGTGCCTACAA 651

QY 1230 GAGGGATATCTGTATCAAACTTGCTTAA 1260
 Db 652 AAGGGATATCTGTATCAAACTTGCTTAA 682

FEATURES

source

Location/Qualifiers

RESULT 13
 B1333581
 LOCUS
 DEFINITION
 mRNA sequence.
 B1333581
 ACCESSION
 B1333581
 VERSION
 B1333581.1 GI:15018238
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC <http://mgs.nci.nih.gov/>.
 1 (bases 1 to 682)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLCMI342 row: 3 column: 17
 High quality sequence stop: 674.

ORIGIN

source

Location/Qualifiers

1..682
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5139184"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

ORIGIN

Query Match

46.3%; Score 583.6; DB 4; Length 682;

Best Local Similarity 97.1%; Pred. No. 1.5e-158;

Matches 647; Conservative 0; Mismatches 14; Indels 5; Gaps 5;

QY 1 ATGCGCGGATCCGCGCTGCGCTGCGGCGCGAGCGCGCGCTTTCTGAGGCGCTTCGGG 60
 Db 19 ATGCGCGGATCCGCGCTGCGCTGCGGCGCGAGCGCGCGCTTTCTGAGGCGCTTCGGG 78
 QY 61 CGGCTGTGCGAGTACAGAGCGCTGTGGGTAGCGGCTCTCCGCTCGGTGTATGCGGTT 120
 Db 79 CGGCTGTGCGAGTACAGAGCGCTGTGGGTAGCGGCTCTCCGCTCGGTGTATGCGGTT 138
 QY 121 CGCTGTGCGGGAACCTGCTGCTGCGCGCGCGCGCGCTCAGGCACTTCTTGCGCGCAGA 180
 Db 139 CGCTGTGCGGGAACCTGCTGCTGCGCGCGCGCGCGCTCAGGCACTTCTTGCGCGCAGA 198
 QY 181 ACCACCGGGGCTGCGGCTCTGCTGCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
 Db 199 ACCACCGGGGCTGCGGCTCTGCTGCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 258
 QY 241 GAACAGTTGCAAGGTCACAGAAACATCTGACTTTGTATGAGAGTGTTCATCCACTTT 300
 Db 259 GAACAGTTGCAAGGTCACAGAAACATCTGACTTTGTATGAGAGTGTTCATCCACTTT 318
 QY 301 TCTCCAAATGTGCATCAAGCTGTCTGTTGAACTCTCGAATGTCAAGTGTTCGGA 360


```

Db      319  TCTCCAAATGTGCATACGCTGTCTGTGCTTGAAGTCTGTGATGTCAGTGTTCGAAA 378
Qy      361  TTGCTCTTATATTCAGTACGAGGCTGTTCATGTGATGTATACAGATTTGGCCCA 420
Db      379  TTGCTCTTATATTCAGTACGAGGCTGTTCATGTGATGTATACAGATTTGGCCCA 438
Qy      421  GATGTTTGGAGGCCCTTGTCTTTCATCATGAGGCTATGTCCATGCGACCTCAA 480
Db      439  GATG-TTTGGAGGCCCTTGTCTTTCATCATGAGGCTATGTCCATGCGACCTCAA 497
Qy      481  CCAAGTAACTATTTGGAGTGCAGAGATGAATGTTTAACTCATTTGAGCTT 540
Db      498  CCAAGTAACTATTTGGAGTGCAGAGATGAATGTTTAACTCATTTGAGCTT 557
Qy      541  AGCTTAAAGAGGCAATCA-AGATGTAAATATATTACAGACAG-ACGGGTATCGGGCTC 598
Db      558  AGCTTAAAGAGGCAATCAAGATGTAAATATATTACAGACAGACGGGTATCGGGCTC 617
Qy      599  CAGAGCAGAAATTCGAAATTTGCTTGGCCAGGCTGCTGCAGAGTGATA-CAGATGT 657
Db      618  CAGAGCAGAAATTCGAAATTTGCTTGG-CCAGGCTGCTGCAGAGTGATACAGAAATGT 676
Qy      658  ACCTCA 663
Db      677  ACCTCA 682

```

```

RESULT 14
BQ225071 888 bp mRNA linear EST 02-MAY-2002
LOCUS     AGENCOURT 7323233 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6016796
DEFINITION 5', mRNA sequence.
ACCESSION BQ225071
VERSION    BQ225071.1 GI:20406471
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 888)
            NIH-MGC http://mgi.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: LHM13215 row: a column: 21
            High quality sequence stop: 581.
            Location/Qualifiers
                1..888
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6016796"
                /tissue_type="epithelioid carcinoma"
                /lab_host="DH10B (phage-resistant)"
                /clone_1ib="NIH MGC 70"
                /note="Organ: pancreas; Vector: pCMV-SPORT6; Site: 1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.1 Kb. Library constructed by Life
                Technologies."

```

```

FEATURES
    source
ORIGIN
Query Match 45.9%; Score 578; DB 5; Length 888;
Best Local Similarity 97.8%; Pred. No. 7.2e-157;
Matches 618; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

```

```

Qy      1  ATGGCGGAGTCCGGCTGCGCTGGGGCGGAGCGCGCGTTCCTGGAGGCTTTCGGG 60
Db      149  ATGGCGGAGTCCGGCTGCGCTGGGGCGGAGCGCGCGCTTCCTGGAGGCTTTCGGG 208
Qy      61  CGGCTGTGGCAGATACAGAGCCGTCTGGGTACGGCTCTCCGGCTCGGTATATGGGTT 120
Db      209  CGGCTGTGGCAGATACAGAGCCGTCTGGGTACGGCTCTCCGGCTCGGTATATGGGTT 268
Qy      121  CGCTGTGGCGGACCCCTGCTGCGCCCCCGGCGGCTTCAGGACATTTTTCGCGCAGGA 180
Db      269  CGCTGTGGCGGACCCCTGCTGCGCCCCCGGCGGCTTCAGGACATTTTTCGCGCAGGA 328
Qy      181  ACCACCGGGGCTGCGGCTCTGCGCGGAGTATGTTTCGGAAGAGAGGGCGGCGCTG 240
Db      329  ACCACCGGGGCTGCGGCTCTGCGCGGAGTATGTTTCGGAAGAGAGGGCGGCGCTG 388
Qy      241  GAACAGTTGCGAGGTTCAGAAACATCGTACCTTTGTATGAGAGTGTTCATTCACATTT 300
Db      389  GAACAGTTGCGAGGTTCAGAAACATCGTACCTTTGTATGAGAGTGTTCATTCACATTT 448
Qy      301  TCTCCAAATGTGCATACGCTGTCTGTGCTTGAATCTGATGTACGTGTTCGAGNA 360
Db      449  TCTCCAAATGTGCATACGCTGTCTGTGCTTGAATCTGATGTACGTGTTCGAGNA 508
Qy      361  TTGCTCTTATATTCAGTACAGGAGTGTTCATGTGATATACAGCATTTGCGCCGA 420
Db      509  TTGCTCTTATATTCAGTACAGGAGTGTTCATGTGATATACAGCATTTGCGCCGA 568
Qy      421  GATGTTTGGAGG-CCCTTGTCTTTCTTATCATGAGGCTATGTTCATGCGGACTCAA 479
Db      569  GATGTTTGGAGGCCCCCTTGTCTTTCTTATCATGAGGCTATGTTCATGCGGACTCAA 628
Qy      480  ACCACGTAACATATTTGTGAGTGCAGAGATGAATGTTTAACTCATTCATTCAGT 539
Db      629  ACCACGTAACATATTTGTGAGTGCAGAGATGAATGTTTAACTCATTCATTCAGT 688
Qy      540  TAGCTTCAAGAA-GGCAATCAGATGTAAAGTATATTCAGACAGCGGT--ATCGGCG 596
Db      689  TAGCTTCAAGAAAGGAATCAAGATGTAAAGTATATTCAGACAGCGGTATTCGCGGCG 748
Qy      597  TCCAGAGCAGAAATTCGAAATTTGCTTGCGCC 628
Db      749  TCCAGAGCAGAAATTCGAAATTTGCTTGCGCC 780

```

```

RESULT 15
COJ37403 602 bp mRNA linear EST 29-JUL-2004
LOCUS     S1702c17116f1 squirrel testis library 1 Spermophilus lateralis
DEFINITION cDNA clone 17116 5', mRNA sequence.
ACCESSION COJ37403
VERSION    COJ37403.1 GI:50824673
KEYWORDS   EST.
SOURCE      Spermophilus lateralis (golden-mantled ground squirrel)
ORGANISM    Spermophilus lateralis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
            Spermophilus.
            1 (bases 1 to 602)
            Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
            Rogers,J. and Cossins,A.R.
            Microarray analysis of transcriptional changes during hibernation
            in the golden mantled ground squirrel, Spermophilus lateralis
            Unpublished (2004)
            Contact: Andrew R. Cossins
            Laboratory for Environmental Gene Regulation
            University of Liverpool
            School of Biological Sciences, The Biosciences Building, Crown
            Street, Liverpool, United Kingdom, L69 7ZB
            Tel: +44(0)151-795-4510
            Fax: +44(0)151-795-4431
            Email: cossins@liv.ac.uk

```


Vector has been trimmed from this EST.
Plate: 17, row: i, column: 16
Seq primer: pf1c T7 (5'-ATAGCACTCACTATAGGG-3')
High quality sequence stop: 602.
Location/Qualifiers

FEATURES

source

1. 602

/organism="Spermophilus lateralis"
/mol_type="mRNA"
/db_xref="taxon:76772"
/clone="17116"
/sex="Male & female"
/tissue_type="Testis"
/dev_stage="Adult"
/lab_host="E.coli Biotromax DH10B"
/note="Vector: pFLC; Site_1: Sail GTCGAG; Site_2: BamHI
GGATCC; Normalized and subtracted cDNA library prepared
from testis of hibernating and summer animals"

ORIGIN

Query Match 44.4%; Score 560; DB 7; Length 602;

Best Local Similarity 95.8%; Pred. No. 1,1e-151;

Matches 575; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
QY 624 GGGCCAGGCTGGCTGCAGAGTATACAGATGTAACCTCAGCTGTTGATCTGTGAGGCT 683
Db 3 GGGCCAGGCTGGCTGCAGAGTATACAGATGTAACCTCAGCTGTTGATCTGTGAGGCT 62
QY 684 AGGAATCATTTTACTGGAAATGTTCTCAGAAATGAACTGAAACATACAGTCAGATCTCA 743
Db 63 AGGAATCATTTTACTGGAAATGTTCTCAGAAATGAACTGAAACATACAGTCAGATCTCA 122
QY 744 GGAATGGAAGGCAAAACAGTCTGCTATTTATTTGATCAGATTTTCCAGTAAAGCAGTGT 803
Db 123 GGAATGGAAGGCAAAATAGTCTCTATTTATTTGATCAGATTTTCCAGTAAAGCAGTGT 182
QY 804 GAATGCCGCAATTCAGCCTATACCTTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 863
Db 183 GAATGCCGCAATTCAGCCTATACCTTAAGAGACCTTATCAAAAGCATGCTTCATGATGA 242
QY 864 TCCAGCAGAAAGATTCTGCTGAAATGGCATTGTGCAGCCCATTTCTTAAATTCCTTT 923
Db 243 TCCAGCAGAAAGATTCTGCTGAAATGGCATTGTGCAGCCCATTTCTTAAATTCCTTT 302
QY 924 TGGCCCTCATTTGAAATGATCTGTCATGCTTCCCATCCAGTGTAAAGATGCTGAATGT 983
Db 303 TGGCCCTCATTTGAAATGATCTGTCATGCTTCCCATCCAGTGTAAAGATGCTGAATGT 362
QY 984 GCTGATGATGATTTATCTTGGATGAAGAAGATATGAATGTTGTAGAAATGTAA 1043
Db 363 CTTGATGATGATTTATCTTGGATGAAGAAGATATGAATGTTGTAGAAATGTAA 422
QY 1044 AGAGAGATGTCAAAAATATGACCAAGTGTATCTTACTTTTCCAAAGGAAATCCTGG 1103
Db 423 AGAGAGATGTCAAAAATATGACCAAGTGTATCTTACTTTTCCAAAGGAAATCCTGG 482
QY 1104 CAGAGCAAAAGTCTTTTGTGATGATGCAATGCTGTGATTTCCAAAGCTGCGCAAAAT 1163
Db 483 CAGAGGCGCAAGTCTTTTGTGATGATGCAATGCTGTGATTTCCAAAGCTGCGCAAAAT 542
QY 1164 ACTGACTGAAAGATGTTTGTGAGGAAGTTGTTGTGGCTACATTTACCCGCTGAGTGC 1223
Db 543 ACTTACTGAAAGATGTTTGTGAGGAATTCGTTGTGGCTACATTTACCCGCTGAGTGC 602
```

Search completed: November 29, 2004, 17:46:26
Job time : 5313 secs

This Page Blank (uspto)

181 ACCACCGGGGCTGGCCCTCTGCGCGAGTATGTTTCGCAAAAGAGGGGCGCTG 240
241 GAACAGTTGAGGGTCAAGAAACATGTAATCTTTGATGAGTGTTCATCCACTTT 300
241 GAACAGTTGAGGGTCAAGAAACATGTAATCTTTGATGAGTGTTCATCCACTTT 300
301 TCTCCAAATGTGCGATCAGCGTCTGTGTTGAACTCTGAGATGTCAGTGTTCGAA 360
301 TCTCCAAATGTGCGATCAGCGTCTGTGTTGAACTCTGAGATGTCAGTGTTCGAA 360
361 TTGCTCTTATTTCCAGTCAACGAGGTGTTCATGTTGATGATTAACGATTTGCGGCA 420
361 TTGCTCTTATTTCCAGTCAACGAGGTGTTCATGTTGATGATTAACGATTTGCGGCA 420
421 GATGTTTGGAGGCCCTGCTTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 480
421 GATGTTTGGAGGCCCTGCTTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 480
481 CCAAGTACATATTTGTGAGTGCAGAGATGAATGTTTAACTCATATGACTTTGAACTT 540
481 CCAAGTACATATTTGTGAGTGCAGAGATGAATGTTTAACTCATATGACTTTGAACTT 540
541 AGCTTCAAGAGGCAATCAGATGTAATGTAATTTCAACACAGCGGATTCGGGCTCCA 600
541 AGCTTCAAGAGGCAATCAGATGTAATGTAATTTCAACACAGCGGATTCGGGCTCCA 600
601 GAAGCAGAAATTTGCAAAATTTGCTGGCCAGGCTGCGCTGCAGAGTATACAAATGTACC 660
601 GAAGCAGAAATTTGCAAAATTTGCTGGCCAGGCTGCGCTGCAGAGTATACAAATGTACC 660
661 TCAGCTGTGATCTGTGAGGCTTGAAGATCATTTTCTGAAATGTTTCTCAGAAATGAA 720
661 TCAGCTGTGATCTGTGAGGCTTGAAGATCATTTTCTGAAATGTTTCTCAGAAATGAA 720
721 CTGAAACATTCAGATCTCAGAAATGGAAGCAACAGTTCTGTATTTATGATCAC 780
721 CTGAAACATTCAGATCTCAGAAATGGAAGCAACAGTTCTGTATTTATGATCAC 780
781 ATATTTGGCAGTAAAGCAGTGTGAATGCGCAATTCAGCTATCACTTAAGAGACTT 840
781 ATATTTGGCAGTAAAGCAGTGTGAATGCGCAATTCAGCTATCACTTAAGAGACTT 840
841 ATCAAAAGCATGCTTCATGATGATCCAAAGCAGAAATTCCTGCTGAATGAGCATTTGTC 900
841 ATCAAAAGCATGCTTCATGATGATCCAAAGCAGAAATTCCTGCTGAATGAGCATTTGTC 900
901 AGCCCATTTCTTATGATTTCTTTTGGCCCTCATATTTGAAGATCTGTCATGCTTCCACT 960
901 AGCCCATTTCTTATGATTTCTTTTGGCCCTCATATTTGAAGATCTGTCATGCTTCCACT 960
961 CCAAGTCTAAGATGCTGATGATGCTGAGATGATGATTTCTTGGAAATGAAGAGAAAT 1020
961 CCAAGTCTAAGATGCTGATGATGCTGAGATGATGATTTCTTGGAAATGAAGAGAAAT 1020
1021 GAAGATGTTGTAAGATGTAAGAGAGAGATGTCAAAATATATGAGCAAGTGTATCTTA 1080
1021 GAAGATGTTGTAAGATGTAAGAGAGAGATGTCAAAATATATGAGCAAGTGTATCTTA 1080
1081 CTTGTTTCAAAAGAAAAATCTGTGCAAGAGCAAGTCTTTGTGATGCAAAATGCTGCT 1140
1081 CTTGTTTCAAAAGAAAAATCTGTGCAAGAGCAAGTCTTTGTGATGCAAAATGCTGCT 1140
1141 GATTTCCAAAGCTGCGCAAAATTAATGATGAGAGATGTTTGAATGAGAGATTTGTTG 1200
1141 GATTTCCAAAGCTGCGCAAAATTAATGATGAGAGATGTTTGAATGAGAGATTTGTTG 1200
1201 GCTACATTTCAACCGCTGAGTGTCTACAAAGAGGAGATCTGTATCAAACTTGTCTTA 1260
1201 GCTACATTTCAACCGCTGAGTGTCTACAAAGAGGAGATCTGTATCAAACTTGTCTTA 1260

AX262521
LOCUS AX262521 1260 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 12 from Patent WO0173050.
ACCESSION AX262521
VERSION AX262521.1 GI:16511408
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Meyers, R.
TITLE 3714, 16742, 23546, and 13887 novel protein kinase molecules and
uses therefor
JOURNAL Patent: WO 0173050-A 12 04-OCT-2001;
Milemnum Pharmaceuticals, Inc. (US)
FEATURES
source 1..1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 1255.2; DB 6; Length 1260;
Best Local Similarity 99.8%; Pred. No. 1.2e-310; Mismatches 0; Gaps 0;
Matches 1257; Conservative 0; Indels 3; Gaps 0;
QY 1 ATGACCGGATCCGGCTGCGCTGCGGCGGAGCGCGCGCTTTCTGAGGCTTCGCG 60
DB 1 ATGACCGGATCCGGCTGCGCTGCGGCGGAGCGCGCGCTTTCTGAGGCTTCGCG 60
QY 61 CCGCTGTGAGGAGTACAGAGCCGTCTGAGTACCGGCTCTCCGCTCGGTATGCGGT 120
DB 61 CCGCTGTGAGGAGTACAGAGCCGTCTGAGTACCGGCTCTCCGCTCGGTATGCGGT 120
QY 121 CCGTGTGAGGAGTACCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 121 CCGTGTGAGGAGTACCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 ACCACCGGAGCTGCGGCTCTGCGGCGGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
DB 181 ACCACCGGAGCTGCGGCTCTGCGGCGGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
QY 241 GAACAGTTGAGGCTGCAAGAAACATGCTGTTGATGAGTGTTCATCAACTTCACTTT 300
DB 241 GAACAGTTGAGGCTGCAAGAAACATGCTGTTGATGAGTGTTCATCAACTTCACTTT 300
QY 301 TCTCCAAATGTGCGATCAGCGTCTGTGTTGAACTCTGAGATGTCAGTGTTCGAA 360
DB 301 TCTCCAAATGTGCGATCAGCGTCTGTGTTGAACTCTGAGATGTCAGTGTTCGAA 360
QY 361 TTGCTCTTATTTCCAGTCAACGAGGTGTTCATGTTGATGATTAACGATTTGCGGCA 420
DB 361 TTGCTCTTATTTCCAGTCAACGAGGTGTTCATGTTGATGATTAACGATTTGCGGCA 420
QY 421 GATGTTTGGAGGCCCTGCTTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 480
DB 421 GATGTTTGGAGGCCCTGCTTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 480
QY 481 CCAAGTACATATTTGTGAGTGCAGAGATGAATGTTTAACTCATATGACTTTGAACTT 540
DB 481 CCAAGTACATATTTGTGAGTGCAGAGATGAATGTTTAACTCATATGACTTTGAACTT 540
QY 541 AGCTTCAAGAGGCAATCAGATGTAATGTAATTTCAACACAGCGGATTCGGGCTCCA 600
DB 541 AGCTTCAAGAGGCAATCAGATGTAATGTAATTTCAACACAGCGGATTCGGGCTCCA 600
QY 601 GAAGCAGAAATTTGCAAAATTTGCTGGCCAGGCTGCGCTGCAGAGTATACAAATGTACC 660
DB 601 GAAGCAGAAATTTGCAAAATTTGCTGGCCAGGCTGCGCTGCAGAGTATACAAATGTACC 660
QY 661 TCAGCTGTGATCTGTGAGGCTTGAAGATCATTTTCTGAAATGTTTCTCAGAAATGAA 720
DB 661 TCAGCTGTGATCTGTGAGGCTTGAAGATCATTTTCTGAAATGTTTCTCAGAAATGAA 720

Db 661 TCAGCTGTTGATCTGAGCCCTAGGAATCATTTTATCTGGAATGTTCTCAGGAATGAAA 720
QY 721 CTGAACATACAGTCAATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATTTATGATCAC 780
Db 721 CTGAACATACAGTCAATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATTTATGATCAC 780
QY 781 ATATTTCCAGTAAAGCAGTGTGAATGCGCAATTCAGCCTTATCACTTAAGAGACCTT 840
Db 781 ATATTTCCAGTAAAGCAGTGTGAATGCGCAATTCAGCCTTATCACTTAAGAGACCTT 840
QY 841 ATCAAAAGCATGCTTATGATGATCCAGCAGAAAGATTCCTGCTGAATGCGATTGTGC 900
Db 841 ATCAAAAGCATGCTTATGATGATCCAGCAGAAAGATTCCTGCTGAATGCGATTGTGC 900
QY 901 AGCCCATTTCTTATGATTCCTTTTGGCCCTCATATTTGAAGATCTGGCATGCTCCCACT 960
Db 901 AGCCCATTTCTTATGATTCCTTTTGGCCCTCATATTTGAAGATCTGGCATGCTCCCACT 960
QY 961 CCAGTCTAAGCATGCTGATGATGCTGATGATGATTTATCTTGGAAATGAAGAAATAT 1020
Db 961 CCAGTCTAAGCATGCTGATGATGCTGATGATGATTTATCTTGGAAATGAAGAAATAT 1020
QY 1021 GAAGATGTTGATGAAGATGTAAGAGAGAGTCAAAATATGACCAAGTGTATCTCTA 1080
Db 1021 GAAGATGTTGATGAAGATGTAAGAGAGAGTCAAAATATGACCAAGTGTATCTCTA 1080
QY 1081 CTGTGTCGAAGAAATCCTGGCAGAGCAAGTCTTTGTTGATGATGCAAAATGCTGGT 1140
Db 1081 CTGTGTCGAAGAAATCCTGGCAGAGCAAGTCTTTGTTGATGATGCAAAATGCTGGT 1140
QY 1141 GATTCGAAGCTGCGAGAAATTAATGATGATGATGATGATGATGATGATGATGATG 1200
Db 1141 GATTCGAAGCTGCGAGAAATTAATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 GCTACATTTCTACCCGCTGAGTGCCTACAGAGGGGATATCTGATCAAACTTGTCTTA 1260
Db 1201 GCTACATTTCTACCCGCTGAGTGCCTACAGAGGGGATATCTGATCAAACTTGTCTTA 1260

RESULT 3
AX262519 2598 bp DNA linear PAT 26-OCT-2001
LOCUS AX262519
DEFINITION Sequence 10 from Patent WO0173050.
ACCESSION AX262519
VERSION AX262519.1 GI:16511406
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Meyers, R.
TITLE 3714, 16742, 23546, and 13887 novel protein kinase molecules and
uses thereof
JOURNAL Patent: WO 0173050-A 10 04-OCT-2001;
Milleium Pharmaceuticals, Inc. (US)
FEATURES
Source
CDS
1..2598
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
268..1527
/note="unamed protein product"
/codon_start=1
/protein_id="CAD10238.1"
/db_xref="GI:16511407"
/translation="MAGSGCAGWGAEPFRFLFAFRLLMOVSRLLSGSSASVYVRCCG
NPSPPGALKQPLPGTTGAASAAYGPKREBALPOLQGHNITIVIGVFTIHSP
NYSRCLLELDIVSYSEBLLYSHOGCSMMNIOHCRDYBLEALILHREGYHDLK
PRLIWSAENECFKLIDFGLSFKEGNQDVYTIQTDGRAPAEALQCLQVNAQOSTE
CTSAVDLWSLGIILWFSGMKLKHVRSQEMKANSASIIIDHIFAKAVANAGIIPYH
LRLLIKSMLEHDDPSRRIIPAMALCPFSIPFAPHIEDLVMLPTPVILRLINLVDDYL
ENEEYEDVVEDYKECKYKYPVSLVLPKPNRGQVFVEVYVNAAGDSKAAQLTGR

ORIGIN
MEDKENVATFYPPLSAYRRXYLTLL"
Query Match 99.6%; Score 1255.2; DB 6; Length 2598;
Best Local Similarity 99.8%; Pred. No. 1.4e-310;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCGGATCCGGCTGGCGCTGGGCGCGAGCCGCGCTTTCTGAGGCGCTTCGG 60
Db 268 ATGGCGGATCCGGCTGGCGCTGGGCGCGAGCCGCGCTTTCTGAGGCGCTTCGG 327
QY 61 CGGCTGTGGCAGTACAGAGCCGCTGGGTAGCGGCTCTCGCTCGGTATGCGGT 120
Db 328 CGGCTGTGGCAGTACAGAGCCGCTGGGTAGCGGCTCTCGCTCGGTATGCGGT 387
QY 121 CGGCTGTGGCAGACCTGGGCTGGGCGCGGCGGCGCTCAGGAGTTCTTGGCGCAG 180
Db 388 CGGCTGTGGCAGACCTGGGCTGGGCGCGGCGGCGCTCAGGAGTTCTTGGCGCAG 447
QY 181 ACCACCGGGGCTGGGCGCTGGGCGCGGAGTATGTTCCGCAAGAGAGGCGGCGCTG 240
Db 448 ACCACCGGGGCTGGGCGCTGGGCGCGGAGTATGTTCCGCAAGAGAGGCGGCGCTG 507
QY 241 GAACAGTTGACAGGTTCACAGAAACATCGTACTTGTATGAGTGTTCATTCACCTT 300
Db 508 GAACAGTTGACAGGTTCACAGAAACATCGTACTTGTATGAGTGTTCATTCACCTT 567
QY 301 TCTCCAAATGTCCTACAGCTGTCTGTGCTTGAATCTCGGATGTCAGTGTTCGGA 360
Db 568 TCTCCAAATGTCCTACAGCTGTCTGTGCTTGAATCTCGGATGTCAGTGTTCGGA 627
QY 361 TTGCTCTTATTTCCAGTACACAGGATTTGTCATGAGATGATACAGATTGCCCGA 420
Db 628 TTGCTCTTATTTCCAGTACACAGGATTTGTCATGAGATGATACAGATTGCCCGA 687
QY 421 GATGTTTGGAGCCCTTGTCTTTCATCATGAGGCTATGTCATGCGGACCTCAAA 480
Db 688 GATGTTTGGAGCCCTTGTCTTTCATCATGAGGCTATGTCATGCGGACCTCAAA 747
QY 481 CCAGCTAATATTTGTGAGTGCAGAGATGAATGTTTAACTCATGACTTTGAGACTT 540
Db 748 CCAGCTAATATTTGTGAGTGCAGAGATGAATGTTTAACTCATGACTTTGAGACTT 807
QY 541 AGCTTCAAGAGGCAATCAGATGTAAAGTATTTACAGACAGAGGATGCGGCTCA 600
Db 808 AGCTTCAAGAGGCAATCAGATGTAAAGTATTTACAGACAGAGGATGCGGCTCA 867
QY 601 GAAGCAGATTGCAAAATTTGCTTGGCCAGGCTGGCTGAGATGATCAGAAATGTACC 660
Db 868 GAAGCAGATTGCAAAATTTGCTTGGCCAGGCTGGCTGAGATGATCAGAAATGTACC 927
QY 661 TCAGCTGTTGATCTGAGGCTGAGGATCATTTTATCTGGAATGTTCTCAGGAATGAAA 720
Db 928 TCAGCTGTTGATCTGAGGCTGAGGATCATTTTATCTGGAATGTTCTCAGGAATGAAA 987
QY 721 CTGAACATACAGTCAATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATTTATGATCAC 780
Db 988 CTGAACATACAGTCAATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATTTATGATCAC 1047
QY 781 ATATTTCCAGTAAAGCAGTGTGAATGCGCAATTCAGCCTTATCACTTAAGAGACCTT 840
Db 1048 ATATTTCCAGTAAAGCAGTGTGAATGCGCAATTCAGCCTTATCACTTAAGAGACCTT 1107
QY 841 ATCAAAAGCATGCTTATGATGATCCAGCAGAAAGATTCCTGCTGAATGCGATTGTGC 900
Db 1108 ATCAAAAGCATGCTTATGATGATCCAGCAGAAAGATTCCTGCTGAATGCGATTGTGC 1167
QY 901 AGCCCATTTCTTATGATTCCTTTTGGCCCTCATATTTGAAGATCTGGCATGCTCCCACT 960
Db 1168 AGCCCATTTCTTATGATTCCTTTTGGCCCTCATATTTGAAGATCTGGCATGCTCCCACT 1227
QY 961 CCAGTCTAAGCATGCTGATGATGCTGATGATGATTTATCTTGGAAATGAAGAAATAT 1020

Db 1228 CCAGTGTCTGAACTGCTGTAATGTCTGATGATGATTAATCTTGAGATGAGATGAGAAAT 1287
Qy 1021 GAAGATGTTGTAGAAAGATGTAAAAAGAGAGAGTGTCAAAAAATATGATGACCAAGTGTATCTTA 1080
Db 1288 GAAGATGTTGTAGAAAGATGTAAAAAGAGAGAGTGTCAAAAAATATGATGACCAAGTGTATCTTA 1347
Qy 1081 CTTGTTCCAAAGAAAATCTGTGCAAGAGCAAGTCTTTGTTGATGATGCAATGCTGTGGT 1140
Db 1348 CTTGTTCCAAAGAAAATCTGTGCAAGAGCAAGTCTTTGTTGATGATGCAATGCTGTGGT 1407
Qy 1141 GATTCGAAGCTCCGCAAGAAATTAATGATCTGGAAGATGTTTATGATGGAAGTTGTTGGT 1200
Db 1408 GATTCGAAGCTCCGCAAGAAATTAATGATCTGGAAGATGTTTATGATGGAAGTTGTTGGT 1467
Qy 1201 GTACATTTACCCGCTGAGTGCCTACCAAGAGGGATATCTGTATCAAACTTGTCTTAA 1260
Db 1468 GTACATTTACCCGCTGAGTGCCTACCAAGAGGGATATCTGTATCAAACTTGTCTTAA 1527
RESULT 4
LOCUS HSA536197 2909 bp mRNA linear PRI 11-JAN-2003
DEFINITION Homo sapiens mRNA for KIS protein.
ACCESSION AJ536197 GI:27657360
VERSION AJ536197.1 GI:27657360
KEYWORDS KIS protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Bieche, I., Manceau, V., Curmi, P. A., Laurendon, I., Lachkar, S.,
Leroy, K., Vadaud, D., Sobel, A. and Maucuer, A.
TITLE Quantitative RT-PCR reveals a ubiquitous but preferentially neural
expression of the KIS gene in rat and human
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2909)
TITLE Direct Submission
JOURNAL Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du
fer a moulin, 75005 Paris, FRANCE
FEATURES
source
1. .2909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:IMAGP998M213680"
/db_xref="taxon:9606"
/clone="I.M.A.G.E. Consortium CloneID 1455548"
/issue_type="2 pooled tumors (clear cell type)"
1. .2909
/gene="KIS"
/protein_id="CAD60192.1"
/db_xref="GI:27657361"
/translation="MAGSGCAGMBAPBPRLEAFGRULMOYOSRLSGSSASVYRVCCG
NRESPPGALKQPLPBGTTGAASAAATGKRKAALBQJQGRNRYITLYIGVTHIESP
NPFSRLLELDVSVSELLVSSHQSCMMMIQHCARVLEALALHHEGTVADLK
PRLIWSANECEPKLIDFGLSPKBNQDVYIOTDVPRAPEALQCLAQAGVQSDTE
CTSAVDLWSIGIILLEMFSGMKLKTVRSQEWKANSALIDHI PASKAVNNAIPAYH
LRLILSKMLHDDPSRIIPALMLCSPEFSIPFAPHIEDVLMPTPLRLNLVLDLYL
ENREEDVEDVEKRCOKYGPVSLVLEKENGROQVVEYANMGDSKAAQKLTLGR
MDEGRVAVATFPLSAVKRGYLYQTLI"
gene
CDS
ORIGIN
Query Match 99.6%; Score 1255.2; DB 9; Length 2909;
Best Local Similarity 99.8%; Pred. No. 1.4e-310;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGGCGGATCCGCGCTGCGCGCGCGAGAGCCCGCGGCTTTTGTGAGGCGCTTCGAG 60
|||||
Db 1468 GTACATTTACCCGCTGAGTGCCTACCAAGAGGGATATCTGTATCAAACTTGTCTTAA 1527

Db 146 ATGGCGGATCCGCGCTGCGCGCGAGAGCCCGCGGCTTTTGTGAGGCGCTTCGAG 205
Qy 61 CGAGTGTGACAGTACAGAGCCGTCTGAGGTAGCGGCTCTCCGCTCGGTGTATCGGGT 120
Db 206 CGGCTGTGACAGTACAGAGCCGTCTGAGGTAGCGGCTCTCCGCTCGGTGTATCGGGT 265
Qy 121 CGCTGTGCGCAACCTTGAGCTGCGCCCGCGGCGCTTCAGGAGTTCTTGCGCGCAG 180
Db 266 CCTGCTGTGCGCAACCTTGAGCTGCGCCCGCGGCGCTTCAGGAGTTCTTGCGCGCAG 325
Qy 181 ACCACCGGGGCTGCGGCGCTGCGCGCGAGTATGTTTCCGAAAGAGAGCGCGCGCTG 240
Db 326 ACCACCGGGGCTGCGGCGCTGCGCGCGAGTATGTTTCCGAAAGAGAGCGCGCGCTG 385
Qy 241 GAACAGTTGACAGGCTCACAGAAACATCGTATCTTGTATGAGAGTTTACATCCACTT 300
Db 386 GAACAGTTGACAGGCTCACAGAAACATCGTATCTTGTATGAGAGTTTACATCCACTT 445
Qy 301 TCTCCAAATGTCATCAACGCTGTCTGTTGCTTGAACCTCTGAGATGCAATGTTTCG 360
Db 446 TCTCCAAATGTCATCAACGCTGTCTGTTGCTTGAACCTCTGAGATGCAATGTTTCG 505
Qy 361 TTGCTCTTATATTCAGATCACAGAGGTTGTTCCATGATGATATACAGATTTGCGCC 420
Db 506 TTGCTCTTATATTCAGATCACAGAGGTTGTTCCATGATGATATACAGATTTGCGCC 565
Qy 421 GATGTTTGGAGGCGCTGCTTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 480
Db 566 GATGTTTGGAGGCGCTGCTTTCTTCATCATGAGGCTATGTCATGCGGACCTCAAA 625
Qy 481 CCAGCTAACATATGTGAGAGTGCAGAGATGAGATGTTTAACTCATTTGACTTTGACT 540
Db 626 CCAGCTAACATATGTGAGAGTGCAGAGATGAGATGTTTAACTCATTTGACTTTGACT 685
Qy 541 ACCTTCAAGAAAGGAAATCAGATGTAATGATATTCAGACAGAGGATTCGGCTCCA 600
Db 686 ACCTTCAAGAAAGGAAATCAGATGTAATGATATTCAGACAGAGGATTCGGCTCCA 745
Qy 601 GAAGCAGAAATTCAGAAATTTGCTTGCGCCAGGCTGCGAGAGTATACAGATGTACC 660
Db 746 GAAGCAGAAATTCAGAAATTTGCTTGCGCCAGGCTGCGAGAGTATACAGATGTACC 805
Qy 661 TCAGCTGTTGATCTGTGAGAGCTTACGAAATCAATTTTACTGAAATGTTCTCAGAA 720
Db 806 TCAGCTGTTGATCTGTGAGAGCTTACGAAATCAATTTTACTGAAATGTTCTCAGAA 865
Qy 721 CTGAAACATACAGTACAGATCTCAGAAATGAAAGCAAGTTCGCTATTTATGATCAC 780
Db 866 CTGAAACATACAGTACAGATCTCAGAAATGAAAGCAAGTTCGCTATTTATGATCAC 925
Qy 781 ATATTGTCAGTAAAGCAGATGAGTGAATCCGCAATTTCCAGCCTATCAGCTAAGAG 840
Db 926 ATATTGTCAGTAAAGCAGATGAGTGAATCCGCAATTTCCAGCCTATCAGCTAAGAG 985
Qy 841 ATCAAAGAGATGCTTCATGATGATCCAGCAGAAAGATTTCTGCTGAATGCAATTTG 900
Db 986 ATCAAAGAGATGCTTCATGATGATCCAGCAGAAAGATTTCTGCTGAATGCAATTTG 1045
Qy 901 AGCCATTTCTTGAATCTCTTTTCCCTCATATTTGAAGATCTGATATCTTCCACT 960
Db 1046 AGCCATTTCTTGAATCTCTTTTCCCTCATATTTGAAGATCTGATATCTTCCACT 1105
Qy 961 CAGTGTCTAAGCTGCTGATGATGCTGATGATGATTTCTTGGGAATGAAGAGATAT 1020
Db 1106 CAGTGTCTAAGCTGCTGATGATGATGATGATGATTTCTTGGGAATGAAGAGATAT 1165
Qy 1021 GAAGATGTTGTAGAAAGATGTAAAAAGAGAGTGTCAAAAAATATGACCAAGTGTATCTTA 1080
Db 1166 GAAGATGTTGTAGAAAGATGTAAAAAGAGAGTGTCAAAAAATATGACCAAGTGTATCTTA 1225
Qy 1081 CTTGTTCCAAAGAAAATCTGTGCAAGAGCAAGTCTTTGTTGATGATGCAATGCTGTG 1140
Db 1226 CTTGTTCCAAAGAAAATCTGTGCAAGAGCAAGTCTTTGTTGATGATGCAATGCTGTG 1285

QY 1141 GATTCGAAGCTGCGAGAAATTACTGAGAGATGTTGATGAGGAAGTTGTTG 1200
DB 1286 GATTCGAAGCTGCGAGAAATTACTGAGAGATGTTGATGAGGAAGTTGTTG 1345
QY 1201 GCTACATTCTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1260
DB 1346 GCTACATTCTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1405
RESULT 5
CQ412057 4065 bp DNA linear PAT 23-JAN-2004
LOCUS Sequence 19128 from Patent WO0170979.
DEFINITION CQ412057
ACCESSION CQ412057
VERSION CQ412057.1 GI:41319838
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Lee, J. and Lillie, J.
TITLE Genes, compositions, kits, and method for identification,
assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 19128 27-SEP-2001,
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1..4065
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 1255.2; DB 6; Length 4065;
Best Local Similarity 99.8%; Pred. No. 1.5e-310;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCGGAGATCCGGCTGCGCTGGGGCGGAGCCGCGCTTTCGAGGCGCTTCGGG 60
DB 268 ATGGCGGAGATCCGGCTGCGCTGGGGCGGAGCCGCGCTTTCGAGGCGCTTCGGG 327
QY 61 CGGCTGTGCGAGGTACAGAGCCGCTGTGGTAGCGGCTCTCCGCTCGGTATCGGGATT 120
DB 328 CGGCTGTGCGAGGTACAGAGCCGCTGTGGTAGCGGCTCTCCGCTCGGTATCGGGATT 387
QY 121 CGCTGTGCGGCAACCTGTGCTGCGCCCGCGGCGCCCTCAGGCAATTCTTCCGCGAGA 180
DB 388 CGCTGTGCGGCAACCTGTGCTGCGCCCGCGGCGCCCTCAGGCAATTCTTCCGCGAGA 447
QY 181 ACCACCGGGGCTGCGGCTGTGCGCGGAGTATGTTTCCGCAAGAGAGGGCGGCGCTG 240
DB 448 ACCACCGGGGCTGCGGCTGTGCGCGGAGTATGTTTCCGCAAGAGAGGGCGGCGCTG 507
QY 241 GAACAGTTGAGGGGTCAAGAAAATCATGTAATTTGATGAGAGTTTAAATCAATCCATT 300
DB 508 GAACAGTTGAGGGGTCAAGAAAATCATGTAATTTGATGAGAGTTTAAATCAATCCATT 567
QY 301 TCTCCAATGTGCATCAGCTGTCTGTGCTTGAATCTCTGATGTCACTGTCTTGGAA 360
DB 568 TCTCCAATGTGCATCAGCTGTCTGTGCTTGAATCTCTGATGTCACTGTCTTGGAA 627
QY 361 TTGCTCTTATATTCAGTCAACAGGGTGTTCATGTGATGATACAGCATTTGCGCCGA 420
DB 628 TTGCTCTTATATTCAGTCAACAGGGTGTTCATGTGATGATACAGCATTTGCGCCGA 687
QY 421 GATGTTTGAAGGCGCTTGTCTTCTCATCATGAGGGCTATGTCCATGCGGAGCTCAAA 480
DB 688 GATGTTTGAAGGCGCTTGTCTTCTCATCATGAGGGCTATGTCCATGCGGAGCTCAAA 747
QY 481 CCAAGTAACATATTTGAGAGTGACAGAAATGAATGTTTAAATCATTTGAAGCTT 540
DB 748 CCAAGTAACATATTTGAGAGTGACAGAAATGAATGTTTAAATCATTTGAAGCTT 807

QY 541 ACCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGACGGGTATCGGCTCCA 600
DB 808 ACCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGACGGGTATCGGCTCCA 867
QY 601 GAAGCAGAATTCAAAATTTGCTTGGCCAGAGCTGGCTGCGAGAGTATACAGAAATGTAC 660
DB 868 GAAGCAGAATTCAAAATTTGCTTGGCCAGAGCTGGCTGCGAGAGTATACAGAAATGTAC 927
QY 661 TCAGCGTGTATCTGTGAGAGCTAGGAATCAATTTACTGGAATGTCTCAGGAATGAAA 720
DB 928 TCAGCGTGTATCTGTGAGAGCTAGGAATCAATTTACTGGAATGTCTCAGGAATGAAA 987
QY 721 CTGAAAACATACAGTACAGATCTCAGAGATGAGAGCAACAGTTCTGTATTAATTCATAC 780
DB 988 CTGAAAACATACAGTACAGATCTCAGAGATGAGAGCAACAGTTCTGTATTAATTCATAC 1047
QY 781 ATATTTGCCAGTAAAGCAGTGTGAATGCCAATTCAGGCTTATCAGCTTAAGAGACCTT 840
DB 1048 ATATTTGCCAGTAAAGCAGTGTGAATGCCAATTCAGGCTTATCAGCTTAAGAGACCTT 1107
QY 841 ATCAAAAGCATCTTCATGATATCCAGCAGAAAGAAATTCCTGTGAATGSCATTGTGC 900
DB 1108 ATCAAAAGCATCTTCATGATATCCAGCAGAAAGAAATTCCTGTGAATGSCATTGTGC 1167
QY 901 ACCCATTTCTTAGCATTTCTTTTGCCTCCATATTTGAAGATCTGTATGCTTCCCACT 960
DB 1168 ACCCATTTCTTAGCATTTCTTTTGCCTCCATATTTGAAGATCTGTATGCTTCCCACT 1227
QY 961 CCAAGTCTAAGACTGCTGATGTGCTGATGATGATATCTTGGGAATGAAGAGAAATAT 1020
DB 1228 CCAAGTCTAAGACTGCTGATGTGCTGATGATGATATCTTGGGAATGAAGAGAAATAT 1287
QY 1021 GAAGATGTGTGAAGATGTAAAGAGAGTGTCAAAAATATGACAGTGTATCTCTA 1080
DB 1288 GAAGATGTGTGAAGATGTAAAGAGAGTGTCAAAAATATGACAGTGTATCTCTA 1347
QY 1081 CTTGTTCCAAAGAAAATCTGTGCGAGAGCAAGCTTGTGTGATGATGAAATCTGGT 1140
DB 1348 CTTGTTCCAAAGAAAATCTGTGCGAGAGCAAGCTTGTGTGATGATGAAATCTGGT 1407
QY 1201 GCTACATTCTACCCGCTGAGTGCCTTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1260
DB 1468 GCTACATTCTACCCGCTGAGTGCCTTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1527
RESULT 6
AX188296 4065 bp DNA linear PAT 06-AUG-2001
LOCUS AX188296
DEFINITION Sequence 3991 from Patent WO0142467.
ACCESSION AX188296
VERSION AX188296.1 GI:15139769
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Schlegel, R., Deede, J., Berger, A. and Zhao, X.
TITLE Genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 3991 14-JUN-2001,
Millennium Predictive Medicine, Inc. (US)
FEATURES
source 1..4065
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 99.6%; Score 1255.2; DB 6; Length 4065;
 Best Local Similarity 99.8%; Pred. No. 1.5e-310;
 Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 ATGGCGGGAATCCGGCTGCGCTGGGGGCGCGAGACCGCGCGCTTTCTGAGAGGCTTCGGG 60
DB 268 ATGGCGGGAATCCGGCTGCGCTGGGGGCGCGAGACCGCGCGCTTTCTGAGAGGCTTCGGG 327
QY 61 CGGCTGAGGAGTACAGAGCCGCTGAGGAGGCTCTCCGCTCGGATGATCGAGGTT 120
DB 328 CGGCTGAGGAGTACAGAGCCGCTGAGGAGGCTCTCCGCTCGGATGATCGAGGTT 387
QY 121 CGGCTGAGGAGTACAGAGCCGCTGAGGAGGCTCTCCGCTCGGATGATCGAGGTT 180
DB 388 CGGCTGAGGAGTACAGAGCCGCTGAGGAGGCTCTCCGCTCGGATGATCGAGGTT 447
QY 181 ACCACCGGGGCTCGGCTGCTGCGCGAGATGATGTTCCGCAAGAGAGGGCGGCTG 240
DB 448 ACCACCGGGGCTCGGCTGCTGCGCGAGATGATGTTCCGCAAGAGAGGGCGGCTG 507
QY 241 GAACAGTGGAGGGTCAAGAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 508 GAACAGTGGAGGGTCAAGAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
QY 301 TCTCCAAATGTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 568 TCTCCAAATGTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 627
QY 361 TTGCTCTTAATTCAGTCAACGAGGTTGTTCAATGATGATGATGATGATGATGATGATGATGAT 420
DB 628 TTGCTCTTAATTCAGTCAACGAGGTTGTTCAATGATGATGATGATGATGATGATGATGATGAT 687
QY 421 GATGTTTGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 688 GATGTTTGGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747
QY 481 CCAAGTAACTATGAGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 748 CCAAGTAACTATGAGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807
QY 541 AGCTTCAAAAGAGGCAATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 808 AGCTTCAAAAGAGGCAATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
QY 601 GAAGCAAGTGGTGAATGCTGAGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 660
DB 868 GAAGCAAGTGGTGAATGCTGAGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 927
QY 928 TCAAGTGTGATCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 987
DB 928 TCAAGTGTGATCTGAGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 987
QY 721 CTGAAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 780
DB 988 CTGAAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1047
QY 781 ATATTGGCCAGTAAAGAGTGAATGCGCAATTCAGGCTTATCACTAAGAGAGGCTT 840
DB 1048 ATATTGGCCAGTAAAGAGTGAATGCGCAATTCAGGCTTATCACTAAGAGAGGCTT 1107
QY 841 ATCAAAAGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1108 ATCAAAAGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167
QY 901 AGCCCAATCTTTAGCATCTTTTGGCCCTCATATTGAAGATGATGATGATGATGATGATGATGAT 960
DB 1168 AGCCCAATCTTTAGCATCTTTTGGCCCTCATATTGAAGATGATGATGATGATGATGATGATGAT 1227
QY 961 CCAAGTCAAGATGCTGAGTGAATGCTGAGTGAATGCTGAGTGAATGCTGAGTGAATGCTGAGTGA 1020
DB 1228 CCAAGTCAAGATGCTGAGTGAATGCTGAGTGAATGCTGAGTGAATGCTGAGTGAATGCTGAGTGA 1287

```

```

QY 1021 GAAGATGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1288 GAAGATGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347
QY 1081 CTTGTTCCAAAGAAATCTTGGCAGAGAGCAAGCTTGTGATGATGATGATGATGATGATGATGAT 1140
DB 1348 CTTGTTCCAAAGAAATCTTGGCAGAGAGCAAGCTTGTGATGATGATGATGATGATGATGATGAT 1407
QY 1141 GATTCCAAAGCTGCGCAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1408 GATTCCAAAGCTGCGCAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1467
QY 1201 GCTACATTTACCCGCTGAGTCCCTACAGAGGAGATATCTGATCAAACTTGTCTTAA 1260
DB 1468 GCTACATTTACCCGCTGAGTCCCTACAGAGGAGATATCTGATCAAACTTGTCTTAA 1527

```

```

RESULT 7
AXI6543 1260 bp DNA linear PAT 22-JUN-2001
LOCUS AXI6543
DEFINITION Sequence 34 from Patent WO0138503.
ACCESSION AXI6543
VERSION AXI6543.1 GI:14546888
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Ploeman, G.D., Whyte, D., Manning, G.S., Sudarshanam, S.S., Martinez, R.,
TITLE Flanagan, P., and Clary, D.S.
JOURNAL Novel human protein kinases and protein kinase-1-like enzymes
Patent: WO 0138503-A 34 31-MAY-2001;
Sugen, Inc. (US)
FEATURES
source location/Qualifiers
1..1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

```

Query Match 99.5%; Score 1253.6; DB 6; Length 1260;
 Best Local Similarity 99.7%; Pred. No. 3e-310;
 Matches 1256; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 ATGGCGGGAATCCGGCTGCGCTGGGGGCGCGAGACCGCGCGCTTTCTGAGAGGCTTCGGG 60
DB 1 ATGGCGGGAATCCGGCTGCGCTGGGGGCGCGAGACCGCGCGCTTTCTGAGAGGCTTCGGG 60
QY 61 CGGCTGAGGAGTACAGAGCCGCTGAGGAGGCTCTCCGCTCGGATGATCGAGGTT 120
DB 61 CGGCTGAGGAGTACAGAGCCGCTGAGGAGGCTCTCCGCTCGGATGATCGAGGTT 120
QY 121 CGGCTGAGGAGTACAGAGCCGCTGAGGAGGCTCTCCGCTCGGATGATCGAGGTT 180
DB 121 CGGCTGAGGAGTACAGAGCCGCTGAGGAGGCTCTCCGCTCGGATGATCGAGGTT 180
QY 181 ACCACCGGGGCTCGGCTGCTGCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 ACCACCGGGGCTCGGCTGCTGCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 GAACAGTGGAGGGTCAAGAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 GAACAGTGGAGGGTCAAGAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 TCTCCAAATGTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 TCTCCAAATGTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TTGCTCTTAATTCAGTCAACGAGGTTGTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 TTGCTCTTAATTCAGTCAACGAGGTTGTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 420

```


QY 421 GATGTTTGGAGGCCCTTCTTCTTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
 DB 421 GATGTTTGGAGGCCCTTCTTCTTTCATCATGAGGGCTATGTCATGCGGACCTCAAA 480
 QY 481 CCAAGTAAATATGTCAGAGTGAAGAAATGTTTAACTCATGACTTTGACTT 540
 DB 481 CCAAGTAAATATGTCAGAGTGAAGAAATGTTTAACTCATGACTTTGACTT 540
 QY 541 ACCTTCAAGAGGAGGAGTGAAGTGAATGTAATATGTAAGAGAGGAGTGAAGTGAAG 600
 DB 541 ACCTTCAAGAGGAGGAGTGAAGTGAATGTAATGTAAGAGAGGAGTGAAGTGAAG 600
 QY 601 GAAGCAGATTCGAAATATGTCCTGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 660
 DB 601 GAAGCAGATTCGAAATATGTCCTGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 660
 QY 661 TCAGCTGTTGATCTGTCAGAGCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720
 DB 661 TCAGCTGTTGATCTGTCAGAGCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720
 QY 721 CTGAAACATACAGTCAATCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 780
 DB 721 CTGAAACATACAGTCAATCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 780
 QY 781 ATATTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
 DB 781 ATATTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
 QY 841 ATCAAAAGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 DB 841 ATCAAAAGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 QY 901 AGCCCATTTCTTGAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
 DB 901 AGCCCATTTCTTGAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
 QY 961 CCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020
 DB 961 CCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020
 QY 1021 GAAGATGTTGAGAGATGTAAGAGAGAGTGAAGAGAGTGAAGAGAGTGAAGAGAGTGAAGAGAGTGAAG 1080
 DB 1021 GAAGATGTTGAGAGATGTAAGAGAGAGTGAAGAGAGTGAAGAGAGTGAAGAGAGTGAAGAGAGTGAAG 1080
 QY 1081 CTTGTTTCAAG 1140
 DB 1081 CTTGTTTCAAG 1140
 QY 1141 GATTCAG 1200
 DB 1141 GATTCAG 1200
 QY 1201 GCTACATTCCTACCGCTGAGTGCCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 DB 1201 GCTACATTCCTACCGCTGAGTGCCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

RESULT 8
 AX680136 1824 bp DNA linear PAT 29-MAR-2003
 LOCUS AX680136 Sequence 21 from Patent WO0181555.
 DEFINITION AX680136
 ACCESSION AX680136
 VERSION AX680136.1 GI:29369916
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS
 1 Gururajan, R., Paterson, C., Recipon, S. A., Tribouley, C. M.,
 Hafalla, A., Khan, F. Y., Yue, H., Au-Young, J., Bandman, O.,
 Baughn, M. R., Borowsky, M. L., Burford, N., Burkill, J. D., Elliot, V. S.,
 Gandhi, A. R., Kearney, L., Lal, P., Lu, D. A., Lu, Y., Tang, Y. T.,

Zingler, K. A., Griffin, D. A., Hillman, J. J., Marcus, G. A., Nguyen, D. B.,
 Policky, J., Ramkumar, J., Thornton, M., Wallis, N. K. and Walsh, R. T.
 Human Kinases
 Patent: NO 0181555-A 21 01-NOV-2001;
 Incyte Genomics, Inc. (US)
 Location/Qualifiers
 1..1824
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 3107952CB1"

ORIGIN

Query Match 99.5%; Score 1253.6; DB 6; Length 1824;
 Best Local Similarity 99.7%; Pred. No. 3.3e-310;
 Matches 1256; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGCGGAGATCGGCTGCGCTGCGGAGCGGAGCGGCGGCTTCTGAGAGCCCTTGGGG 60
 DB 1 ATGGCGGAGATCGGCTGCGCTGCGGAGCGGAGCGGCGGCTTCTGAGAGCCCTTGGGG 60
 QY 61 CCGCTGTGTCAGAGTCAAGAGCCCTGCGGTAGCGGCTCTCCGCTCGGTGTATGCGGTT 120
 DB 61 CCGCTGTGTCAGAGTCAAGAGCCCTGCGGTAGCGGCTCTCCGCTCGGTGTATGCGGTT 120
 QY 229 CCGCTGTGTCAGAGTCAAGAGCCCTGCGGTAGCGGCTCTCCGCTCGGTGTATGCGGTT 288
 DB 229 CCGCTGTGTCAGAGTCAAGAGCCCTGCGGTAGCGGCTCTCCGCTCGGTGTATGCGGTT 288
 QY 121 CGCTGCTGCGGCAACCTGCGGTGCGGCGGCGGCGGCTCAGAGAGTTCCTGCGGCGGCA 180
 DB 121 CGCTGCTGCGGCAACCTGCGGTGCGGCGGCGGCGGCTCAGAGAGTTCCTGCGGCGGCA 180
 QY 288 CGCTGCTGCGGCAACCTGCGGTGCGGCGGCGGCGGCTCAGAGAGTTCCTGCGGCGGCA 348
 DB 288 CGCTGCTGCGGCAACCTGCGGTGCGGCGGCGGCGGCTCAGAGAGTTCCTGCGGCGGCA 348
 QY 181 ACCACCGGGGCTGCGGCTGCGGCGGCGGAGTATGTTTCCGCAAGAGAGGGCGGCGTG 240
 DB 181 ACCACCGGGGCTGCGGCTGCGGCGGCGGAGTATGTTTCCGCAAGAGAGGGCGGCGTG 240
 QY 349 ACCACCGGGGCTGCGGCTGCGGCGGCGGAGTATGTTTCCGCAAGAGAGGGCGGCGTG 408
 DB 349 ACCACCGGGGCTGCGGCTGCGGCGGCGGAGTATGTTTCCGCAAGAGAGGGCGGCGTG 408
 QY 241 GAACAGTTGCAAGGCTCAAGAAACATCGTACCTTGTATGAGAGTTCATCAATCCACTTT 300
 DB 241 GAACAGTTGCAAGGCTCAAGAAACATCGTACCTTGTATGAGAGTTCATCAATCCACTTT 300
 QY 409 GAACAGTTGCAAGGCTCAAGAAACATCGTACCTTGTATGAGAGTTCATCAATCCACTTT 468
 DB 409 GAACAGTTGCAAGGCTCAAGAAACATCGTACCTTGTATGAGAGTTCATCAATCCACTTT 468
 QY 301 TCTCCAAATGTCGATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 DB 301 TCTCCAAATGTCGATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 QY 468 TCTCCAAATGTCGATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 528
 DB 468 TCTCCAAATGTCGATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 528
 QY 361 TTGCTCTTATATTCAGTCAACAGAGGTTTTCATGATGATGATGATGATGATGATGATGATGATGATG 420
 DB 361 TTGCTCTTATATTCAGTCAACAGAGGTTTTCATGATGATGATGATGATGATGATGATGATGATGATG 420
 QY 528 TTGCTCTTATATTCAGTCAACAGAGGTTTTCATGATGATGATGATGATGATGATGATGATGATGATG 588
 DB 528 TTGCTCTTATATTCAGTCAACAGAGGTTTTCATGATGATGATGATGATGATGATGATGATGATGATG 588
 QY 421 GATGTTTGGAGGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 480
 DB 421 GATGTTTGGAGGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 480
 QY 589 GATGTTTGGAGGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 648
 DB 589 GATGTTTGGAGGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 648
 QY 481 CCAAGTAAATATGTCAGAGTGAAGAAATGTTTAACTCATGACTTTGACTT 540
 DB 481 CCAAGTAAATATGTCAGAGTGAAGAAATGTTTAACTCATGACTTTGACTT 540
 QY 649 CCAAGTAAATATGTCAGAGTGAAGAAATGTTTAACTCATGACTTTGACTT 708
 DB 649 CCAAGTAAATATGTCAGAGTGAAGAAATGTTTAACTCATGACTTTGACTT 708
 QY 541 ACCTTCAAGAGGAGAGATCAGAGTGAAGATATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 DB 541 ACCTTCAAGAGGAGAGATCAGAGTGAAGATATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 QY 709 ACCTTCAAGAGGAGAGATCAGAGTGAAGATATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
 DB 709 ACCTTCAAGAGGAGAGATCAGAGTGAAGATATATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
 QY 601 GAAGCAGATTCGAAATATGTCCTGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660
 DB 601 GAAGCAGATTCGAAATATGTCCTGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660
 QY 769 GAAGCAGATTCGAAATATGTCCTGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 828
 DB 769 GAAGCAGATTCGAAATATGTCCTGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 828
 QY 661 TCAGCTGTTGATCTGTCAGAGCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720
 DB 661 TCAGCTGTTGATCTGTCAGAGCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720
 QY 829 TCAGCTGTTGATCTGTCAGAGCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 888
 DB 829 TCAGCTGTTGATCTGTCAGAGCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 888
 QY 721 CTGAAACATACAGTCAATCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 780
 DB 721 CTGAAACATACAGTCAATCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 780
 QY 889 CTGAAACATACAGTCAATCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 948
 DB 889 CTGAAACATACAGTCAATCTGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 948
 QY 781 ATATTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
 DB 781 ATATTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
 QY 949 ATATTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1008
 DB 949 ATATTGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1008

QY 841 ATCAAAAGCATGCTTCATGATGATCCAGAGAGAAATTCCTGCTGAAAATGGCATTTGTC 900
 DB 1009 ATCAAAAGCATGCTTCATGATGATCCAGAGAGAAATTCCTGCTGAAAATGGCATTTGTC 1068
 QY 901 AGCCCATTTCTTAAGCATTCCTTTTGGCCCTCATATTGAAAGATTTGGTCATGCTTCCACT 960
 DB 1069 AGCCCATTTCTTAAGCATTCCTTTTGGCCCTCATATTGAAAGATTTGGTCATGCTTCCACT 1128
 QY 961 CCAGTCTAAGATGCTGGAATGCTGATGATGATGATTTATGTTGGAAATGAGAAATAT 1020
 DB 1129 CCAGTCTAAGATGCTGGAATGCTGATGATGATGATTTATGTTGGAAATGAGAAATAT 1188
 QY 1021 GAAGATGTTGTAGAGATGTAAAAAGAGATGTCAAAAATATGACACAGTGTATCTCTA 1080
 DB 1189 GAAGATGTTGTAGAGATGTAAAAAGAGATGTCAAAAATATGACACAGTGTATCTCTA 1248
 QY 1081 CTGTGTTCCAAAGGAAAATCTGCGACAGACAGATCTTTTGTGATGATGCAAACTTGTCTTA 1140
 DB 1249 CTGTGTTCCAAAGGAAAATCTGCGACAGACAGATCTTTTGTGATGATGCAAACTTGTCTTA 1308
 QY 1141 GATTCCAAAGCTGCGAGAAATTAATGACTGAGAGAGATGTTGATGGAAGTTGTTGTG 1200
 DB 1309 GATTCCAAAGCTGCGAGAAATTAATGACTGAGAGAGATGTTGATGGAAGTTGTTGTG 1368
 QY 1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTA 1260
 DB 1369 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTA 1428

RESULT 9
 BC058732 2041 bp mRNA linear ROD 30-JUN-2004
 LOCUS BC058732
 DEFINITION Mus musculus kinase interacting with leukemia-associated gene (lethalmn), mRNA (cDNA clone MGC:64170 IMAGE:6414877), complete cds.

ACCESSION BC058732
 VERSION BC058732.1 GI:37194892
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2041)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, K.H., Scheaffer, C.P., Bhat, N.K., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, T.L., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Stachenko, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richard, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schermer, A., Schrein, J.E., Jones, S.J., and Marra, M.A., 2003. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBLISHED 12477932
 REFERENCE 2 (bases 1 to 2041)
 Strausberg, R.
 TITLE Direct Submission
 AUTHORS Submitted (24-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldo, M.F., Akabou, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Koppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Smith, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

FEATURES
 source
 1..2041
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:64170 IMAGE:6414877"
 /tissue_type="Brain, pool of enriched mouse brain tissues from 13.5, 14.5, 16.5 and 17.5 dpc"
 /clone_lib="NIH BMAP_FRO"
 /lab_host="DH10B"
 /note="vector: pYX-ASC"
 1..2041
 /gene="Kist"
 /note="synonym: KIS"
 /db_xref="locusid:16589"
 /db_xref="MGI:1341908"
 154..1413
 /gene="Kist"
 /codon_start=1
 /product="kinase interacting with leukemia-associated gene (lethalmn)"
 /protein_id="AAH58732.1"
 /db_xref="GI:37194893"
 /db_xref="MGI:1341908"
 /translation="MAGSCAGAGAPRFLFAGRLAQVQSLRSGSSASVYRVCCGTPGSPGALKQFLPPTGGAASABYGRKRRLBLOGHRNVTVTGVTTHFSPNPIIWSAENECFLIDFGLFKEGNDVXTIQTGTPAPRPAELONCIAQAGLQDTECTSAVDLSMHDHPGRIRIPAEMLCSPFSPISPPAPRIEDLVMLPTVLLNLVLDYTL ENEDEYEDVEDEVEKCECKYGVVSLVLPKKNPGRGVFEVAYANAGDSKAQKLTGR MFDGRFVATFPLPAPYKRGVLYQTL"

ORIGIN
 Query Match 88.4%; Score 1114.4; DB 10; Length 2041;
 Best Local Similarity 92.8%; Pred. No. 1.7e-274;
 Matches 1169; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ATGGCGGGAATCCGCGCTCGGCGCGCGGAGCCGCGCGCTTTCTGAGGCGCTTCGGG 60
 DB 154 ATGGCGGGAATCCGCGCTCGGCGCGCGGAGCCGCGCGCTTTCTGAGGCGCTTCGGG 213
 QY 61 CGGCTGTGACAGTACAGAGCCGCTGAGGATACGCGCTCTCCGCGCTGAGTATCGGGTT 120
 DB 214 CGGCTGTGACAGTACAGAGCCGCGCTGAGGATACGCGCTCTCCGCGCTGAGTATCGGGTT 273
 QY 121 CGCTGTGCGGACACCTTGCTGCGCCCGCGCGCTTACGACGAGTTCTTCCGCCAGGA 180
 DB 274 CGGTGTGCGGACACCTTGCTGCGCCCGCGCGCTTACGACGAGTTCTTCCGCCAGGA 333
 QY 181 ACCACCGGCGCTGCGGCTCTGCGCGCGGAGTATGTTTCCGAAAGAGGGCGGCGCTG 240

Db 334 ACCACGGGGCTGCGCCCTGCGCCGAGATGTTTCGCAAGAGAGGGCGGCTG 393
 QY 241 GAACAGTTGAGGGGTCAAGAAACATGTCATCTTGTATGAGTGTTCATCACTT 300
 Db 394 GAGCAATTGAGGGGTCAAGAAACATGTCATCTTGTATGAGTGTTCATCACTT 453
 QY 301 TCTCCAAATGTCATCAAGCTGTCTGTGCTTGAATCCCTGAGTGTCAAGTGTTCGGA 360
 Db 454 TCTCCAAATGTCATCAAGCTGTCTGTGCTTGAATCCCTGAGTGTCAAGTGTTCGGA 513
 QY 361 TTGCTCTTATTTCCAGTCAACGAGGTTTTCATGTATGATATACATTCGCCCCA 420
 Db 514 TTGCTCTTATTTCCAGTCAACGAGGTTTTCATGTATGATATACATTCGCCCCA 573
 QY 421 GATGTTTGGAGGCGCTGCTTTCATCATAGAGGCTATGCTCATGCGACTCAAA 480
 Db 574 GATGTTTGGAGGCGCTGCTTTCATCATAGAGGCTATGCTCATGCGACTCAAA 633
 QY 481 CCACTTAACATATTTGAGTGCAGAGAAATGAATGTTTAACTCAATTCATTCGACTT 540
 Db 634 CCACTTAACATCTGTGAGTGCAGAGAAATGAATGTTTAACTCAATTCATTCGACTT 693
 QY 541 AGCTTAAAGAGGCAATCAGATGTAAATATTTCAAGACAGAGGCTATTCGAGCTCCA 600
 Db 694 AGCTTAAAGAGGCAATCAGATGTAAATATTTCAAGACAGAGGCTATTCGAGCTCCA 753
 QY 601 GAAGCAGAAATGCAAAATGCTTGGCCGAGGCTGCGAGTATCAGAAATGATAC 660
 Db 754 GAAGCAGAGTGCAGAAATGCTTGGCCGAGGCTGCGAGTATCAGAAATGATAC 813
 QY 661 TCAGCTGTGATCTGAGAGCTGAGAAATCATTTTACTGGAAATGTTCTCAGAGATGAAA 720
 Db 814 TCAGCTGTGATCTGAGAGCTGAGAAATCATTTTACTGGAAATGTTCTCAGAGATGAAA 873
 QY 721 CTGAAAATATCAGTCAAGATCTCAGAAATGAGGCAAAAGTTCTGCTATTTATTCAC 780
 Db 874 CTGAAAATATCAGTCAAGATCTCAGAAATGAGGCAAAAGTTCTGCTATTTATTCAC 933
 QY 781 ATATTTGCAAGTAAAGAGTGAATGCGCAATTCAGCCATATCACTTAAGAGCTT 840
 Db 934 ATATTTGCAAGTAAAGAGTGAATGCGCAATTCAGCCATATCACTTAAGAGCTT 993
 QY 841 ATCAAAAGCATGCTTATGATGATTCAGAGCAAGAAATTCCTGCTGAAATGCAATTTGTC 900
 Db 994 ATCAAAAGCATGCTTATGATGATTCAGAGCAAGAAATTCCTGCTGAAATGCAATTTGTC 1053
 QY 901 AGCCCATTTCTTGAATCTCTTTTGCCTTCATATTTGAAGATCTGATCTTCCACT 960
 Db 1054 AGCCCATTTCTTGAATCTCTTTTGCCTTCATATTTGAAGATCTGATCTTCCACT 1113
 QY 961 CCAAGTCTAAGATCTGCTGAATGTCGTGATGATTTATTTGGGAATGAAGAGAAAT 1020
 Db 1114 CCAAGTCTAAGATCTGCTGAATGTCGTGATGATTTATTTGGGAATGAAGAGAAAT 1173
 QY 1021 GAAGATGTTAGAGATGTAAAGAGAGATCAAAAATATGACCAAGGATCTCTA 1080
 Db 1174 GAAGATGTTAGAGATGTAAAGAGAGATCAAAAATATGACCAAGGATCTCTA 1233
 QY 1081 CTGTGTTCAAAAGAAATCTGCGAGAGCAAGCTCTTGTGATGATGCAAAATGCTGCT 1140
 Db 1234 CTGTGTTCAAAAGAAATCTGCGAGAGCAAGCTCTTGTGATGATGCAAAATGCTGCT 1293
 QY 1141 GATTCCAAAGCTGCGAGAAATTAATGATCTGGAAGATGTTTGAATGGAAATTTGTTG 1200
 Db 1294 GATTCCAAAGCTGCTCAGAAATGTCGATCTGGAAGATGTTTGAATGGAAATTTGTTG 1353
 QY 1201 GCTACATTTCAACCGCTGAGTGCCTCAAGAGGGGATATCTGTATCAAACTTGTCTTAA 1260
 Db 1354 GCTACATTTCAACCGCTGAGTGCCTCAAGAGGGGATATCTGTATCAAACTTGTCTTAA 1413

RESULT 10
 AC129141_1/c

WPCOMMENT
 Sequence split into 4 fragments LOCUS AC129141 Accession AC129141
 Fragment Name Begin End
 AC129141_0 1 110000
 AC129141_1 100001 210000
 AC129141_2 200001 310000
 AC129141_3 300001 363206
 Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH2)

Query Match 88.2%; Score 1111.2; DB 2; Length 110000;
 Best Local Similarity 92.6%; Pred. No. 2.5e-273;
 Matches 1167; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 1 ATGCGCGGATCCGGCTGCGCCCTGCGCGAGAGCCGCGCTTTCGAGAGCTTTCGG 60
 Db 70316 ATGCGCGGATCCGGCTGCGCCCTGCGCGAGAGCCGCGCTTTCGAGAGCTTTCGG 70257
 QY 61 CGGCTGTGGCAGGTACAGAGCCGTGTGGTAGCGGCTCTCCGCTCGGTGTATCGGTT 120
 Db 70256 CGGCTGTGGCAGGTACAGAGCCGTGTGGTAGCGGCTCTCCGCTCGGTGTATCGGTTG 70197
 QY 121 CGCTGTGGGCAACCTTGCTGCGCCCGGCGGCTTCAAGCAATTTCTGCGCGCAGA 180
 Db 70196 CGCTGTGGGCAATTCAGGCTGCGCCCGGCGGCTTCAAGCAATTTCTGCTCGGGA 70137
 QY 181 ACCACCGGGGCTGCGGCTCTGCGCGAGTATGTTTCCGAAAGAGAGGCGGCTG 240
 Db 70136 ACCACCGGGGCTGCGGCTCTGCGCGAGTATGTTTCCGAAAGAGAGGCGGCTG 70077
 QY 241 GAACAGTTGCAAGGTCACAGAAACATGCTGATGTATGAGTGTTCATCACTT 300
 Db 70076 GAACAGTTGCAAGGTCACAGAAACATGCTGATGTATGAGTGTTCATCACTT 70017
 QY 301 TCTCCAAATGTCATCAAGCTGTCTGTGCTTGAATCTGGAATGATGATTTTCGGA 360
 Db 70016 TCTCCAAATGTCATCAAGCTGTCTGTGCTTGAATCTGGAATGATGATTTTCGGA 69957
 QY 361 TTGCTCTTATTTCCAGTCAACAGGTTTTCATGTGATATCAGATTTGCCCGCA 420
 Db 69956 TTGCTCTTATTTCCAGTCAACAGGTTTTCATGTGATATCAGATTTGCCCGCA 69897
 QY 421 GATGTTTGGAGGCGCTTCTTTCATCAATGAAGGCTATGTCATGCGGACTCAAA 480
 Db 69896 GATGTTTGGAGGCGCTTCTTTCATCAATGAAGGCTATGTCATGCGGACTCAAA 69837
 QY 481 CCACTTAACATTTGAGTGCAGAGAAATGATTTTAACTATGATCTTTCGACTT 540
 Db 69836 CCACTTAACATCTGTGAGTGCAGAGAAATGATTTTAACTATGATCTTTCGACTT 69777
 QY 541 AGCTTCAAAAGAGGCAATCAGAGTAAAGTATTTCAAGAGAGGGATATGCGGCTCA 600
 Db 69776 AGCTTCAAAAGAGGCAATCAGAGTAAAGTATTTCAAGAGAGGGATATGCGGCTCA 69717
 QY 601 GAAGCAATTCGAAATTTGCTTGGCCAGGCTGCGCTGAGAGTATACAGAAATGTA 660
 Db 69716 GAAGCAATTCGAAATTTGCTTGGCCAGGCTGCGCTGAGAGTATACAGAAATGTA 69657
 QY 661 TCAAGCTGTGATCTGAGAGCTTGAAGATCAATTTTACTGGAATGTTCTCAGAAATGAAA 720
 Db 69656 TCAAGCTGTGATCTGAGAGCTTGAAGATCAATTTTACTGGAATGTTCTCAGAAATGAAA 69597
 QY 721 CTGAAAATACAGTCAAGTCTCAGAGATGGAAGAGGCAACAGTTCTGCTATTTATGATAC 780
 Db 69596 CTGAAAATACAGTCAAGTCTCAGAGATGGAAGAGGCAACAGTTCTGCTATTTATGATAC 69537
 QY 781 ATATTTGCAAGTAAAGAGTGTGATGTCGCAATTTCAAGCTTATCACTTAAGAGACTT 840
 Db 69536 ATATTTGCAAGTAAAGAGTGTGATGTCGCAATTTCAAGCTTATCACTTAAGAGACTT 69477
 QY 841 ATCAAAAGCATCTTCAATGATATCAAGCAGAGAAATTTCTGCTGAATGCAATTTGTC 900
 Db 69476 ATCAAAAGCATCTTCAATGATATCAAGCAGAGAAATTTCTGCTGAATGCAATTTGTC 69417

```

QY      901 AGCCCATCTTTAGCATCTCTTTGCCCTCATATATGAGATCTGTCATGCTCCACT 960
Db      69416 AGCCCATCTTTAGCATCTCTTTGCCCTCATATATGAGATCTGTCATGCTCCACT 69357
QY      961 CCGATGCTAGATCTGCTGATGCTGCTGATGATGATATCTTGGGATGAGAGAAATAT 1020
Db      69356 CCGATGCTAGATCTCTCATATGCTGATGATGATGATATCTTGGGATGAGAAATAT 69297
QY      1021 GAAGATGTTAGAGATGTAAGAGAGAGATGTCAAAATAATGAGACCAAGTATCTCTA 1080
Db      69296 GAAGATGTTAGAGATGTAAGAGAGATGTCAAAATAATGAGACCAAGTATCTCTC 69237
QY      1081 CTGTTCAGAAAGAAAATCTGCGACAGACAGATCTTTGTTGAGATGCAAAATGCTGCT 1140
Db      69236 CTGTTCAGAAAGAAAATCTGCGACAGACAGATCTTTGTTGAGATGCAAAATGCTG 69177
QY      1141 GATTCCAAAGCTGCGCAAGAAATTAATGATCTGCAAGATGTTGATGAGAAAGTTGTTGG 1200
Db      69176 GATTCCAAAGCTGCTCAGAAATGCTGACCTGAGAGAGATGTTGATGAGAAAGTTGTTG 69117
QY      1201 GCTACATTTACCGCGTGAAGTGCCTACAGAGAGGAGATATCTGTATCAAACTGCTTTAA 1260
Db      69116 GCTACATTTACCGCGTGAAGTGCCTACAGAGAGGAGATATCTTTATCAAACTGCTTTAA 69057

RESULT 11
RINKSRNA      1633 bp      mRNA      linear      ROD 08-JUL-2002
LOCUS      R.norvegicus mRNA for Kis protein.
ACCESSION      X98374.1 GI:1403531
VERSION      Kis gene.
KEYWORDS      Rattus norvegicus (Norway rat)
SOURCE      Rattus norvegicus
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
      Rattus.
REFERENCE      1
AUTHORS      Maucuer, A., Ozon, S., Manceau, V., Gavet, O., Lawler, S., Curmi, P. and
      Sobel, A.
      Kis is a protein kinase with an RNA recognition motif
      J. Biol. Chem. 272 (37), 23151-23156 (1997)
MEDLINE      97435279
PUBMED      9287318
REFERENCE      2 (bases 1 to 1633)
AUTHORS      Maucuer, A.
      Direct Submission
      Submitted (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer
      Moulin, 75005 Paris, FRANCE
      Location/Qualifiers
      1. 1633
         /organism="Rattus norvegicus"
         /mol_type="mRNA"
         /db_xref="taxon:10116"
         /clone_lib="7"
         /dev_stage="neonate"
         /gene="Kis"
         /gene="KIS"
         /gene="KIS"
         /codon_start=1
         /protein_id="CAA67021.1"
         /db_xref="GI:1403532"
         /db_xref="SwIsr-Pict:063285"
         /db_xref="GDA:063285"
      CDS
      gene
      CDS

```

```

ORIGIN
Query Match      87.8%; Score 1106.4; DB 10; Length 1633;
Best Local Similarity 92.4%; Pred. No. 1.8e-272;
Matches 1164; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

MFDKPVATFVPLSYRRGYLYOTLL"
QY      1 ATGGCGGAGATCCGGCTGCGCCCTGAGGCGCGAGCCGCGCTTTTCTGAGAGCCCTTCCGG 60
Db      39 ATGGCGGAGATCCGGCTGCGCCCTGAGGCGCGAGCCGCGCTTTCTGAGAGCCCTTCCGG 98
QY      61 CGGCTGTGGCAGTACAGAGCCCTGCGGTACCGGCTCTCGGCTCGGTATGCGGT 120
Db      99 CGGCTGTGGCAGTACAGAGCCGCTGAGCCGCTCTCGGCTCGGTATGCGGT 158
QY      121 CGGCTGTGGCAGTACCGTGGCTCGGCCCGGCGGCTCAGGAGATTCTGTCGCGCAGGA 180
Db      159 CGGCTGTGGCAGTACCGTGGCTCGGCCCGGCGGCTCAGGAGATTCTGTCGCGCAGGA 218
QY      181 ACCACCGGAGCTGCGGCTCTGCGCGAGTATGCTTCCGAAAGAGAGGCGGCGCTG 240
Db      219 ACCACCGGAGCTGCGGCTCTGCGCGAGTATGCTTCCGAAAGAGAGGCGGCGCTG 278
QY      241 GAAACGTTGACAGGCTACAGAAACATCGTACTTTGATGAGTGTTAACAATCACTTT 300
Db      279 GAAACGTTGACAGGCTACAGAAACATCGTACTTTGATGAGTGTTAACAATCACTTT 338
QY      301 TCTCCAAATGTCGATCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db      339 TCTCCAAATGTCGATCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 398
QY      361 TTGCTCTTATATTCAGTACAGAGGCTTTCATGATGATGATGATGATGATGATGATGATGATG 420
Db      399 TTGCTCTTATATTCAGTACAGAGGCTTTCATGATGATGATGATGATGATGATGATGATGATG 458
QY      421 GATGTTTGGAGGCGCTTCTTCTTCTATCATGAGGCTATGTCATGCGGAGCTCAAA 480
Db      459 GATGTTTGGAGGCGCTTCTTCTTCTATCATGAGGCTATGTCATGCGGAGCTCAAA 518
QY      481 CCAGTAACATATTTGAGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db      519 CCAGTAACATATTTGAGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGATGATG 578
QY      541 AGCTTCAAAGAGGAGATCAGAGATGTAAGTATATTCAGACAGAGGATCGGCTCCA 600
Db      579 AGCTTCAAAGAGGAGATCAGAGATGTAAGTATATTCAGACAGAGGATCGGCTCCT 638
QY      601 GAAGCAAGTTCAGAAATGCTTGGCCCAAGCTGCGTGCAGAGTATCAGATGATGATGATGATG 660
Db      639 GAAGCAAGTTCAGAAATGCTTGGCCCAAGCTGCGTGCAGAGTATCAGATGATGATGATGATG 698
QY      661 TCAGCGTTGATCTGAGAGGCTTAGAATCATTTTACTGGAATGTTTCTCGAGAAATGAA 720
Db      699 TCAGCGTTGATCTGAGAGGCTTAGAATCATTTTACTGGAATGTTTCTCGAGAAATGAA 758
QY      721 CTGAAACATACAGTACATCTCAGAGATGAAAGGCAACAGTTCTGTATTAATGATGATGATG 780
Db      759 CTGAAACATACAGTACATCTCAGAGATGAAAGGCAACAGTTCTGTATTAATGATGATGATG 818
QY      781 ATATTGGCAGTAAGACAGTGTGATATCGCGCAATTCAGGCTTACCTTAAGAGACTTT 840
Db      819 ATATTGGCAGTAAGACAGTGTGATATCGCGCAATTCAGGCTTACCTTAAGAGACTTT 878
QY      841 ATCAAAAGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Db      879 ATCAAAAGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 938
QY      901 AGCCATCTTTAGCATCTCTTTGCCCTCATATATGAGATCTGTCATGCTCCACT 960
Db      939 AGCCATCTTTAGCATCTCTTTGCCCTCATATATGAGATCTGTCATGCTCCACT 998
QY      961 CCGATGCTAGATCTGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020

```

Db 999 CCAAGTCAGACTCCTCAATGTGCTGGATGATGACTATCTTGAATAATGAAATAT 1058
 QY 1021 GAAGATGTTGTAGAAATGTAAAGAGAGTCAAAAATATGACCAAGTGTATCTTA 1080
 Db 1059 GAAGATGTTGTAGAAATGTAAAGAGAGTCAAAAATATGACCAAGTGTATCTTA 1118
 QY 1081 CTGTTTCAAGAGAAATCTGCGACAGACAAAGTCTTTGTTGATGACCAATGCTGCT 1140
 Db 1119 CTGTTTCAAGAGAAATCTGCGACAGACAAAGTCTTTGTTGATGACCAATGCTGCT 1178
 QY 1141 GATTCAGAGAGTCTGCGACAGACAAAGTCTTTGTTGATGACCAATGCTGCTGCT 1200
 Db 1179 GATTCAGAGAGTCTGCGACAGACAAAGTCTTTGTTGATGACCAATGCTGCTGCT 1238
 QY 1201 GCTACATTCACCGCGTGAAGTCTCAAGAGAGGATATCTGTATCAAACTGCTTTAA 1260
 Db 1239 GCTACATTCACCGCGTGAAGTCTCAAGAGAGGATATCTGTATCAAACTGCTTTAA 1298
 RESULT 12 3244 bp mRNA linear ROD 08-SEP-1999
 RNU70372
 LOCUS Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
 DEFINITION complete cds.
 U70372 GI:5821767
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 3244)
 and Eipper,B.A., Darlington,D.N., Johnson,R.C., Darlington,D.N., Mains,R.E.
 Novel proteins that interact with the COOH-terminal cytosolic
 routing determinants of an integral membrane peptide-processing
 enzyme
 J. Biol. Chem. 271 (45), 28636-28640 (1996)
 JOURNAL MEDLINE 8910496
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
 Determinants of the Peptide Processing Enzyme Peptidylglycine
 alpha-Aminating Monooxygenase
 J. Biol. Chem. 274 (1999) In press
 3 (bases 1 to 3244)
 Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
 and Eipper,B.A.
 Direct Subunit
 Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 4 (bases 1 to 3244)
 Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
 Eipper,B.A. and Mains,R.E.
 Direct Subunit
 Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
 School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
 21205, USA
 REMARK
 COMMENT
 FEATURES
 source
 On Sep 3, 1999 this sequence version replaced gi:1698778.
 Location/Qualifiers
 1..3244
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 162..11421
 /function="serine protein kinase"
 /note="P-CIP2"
 /codon_start=1

/product="PAM COOH-terminal interactor protein 2"
 /protein_id="AAC53031.2"
 /db_xref="GI:5821768"
 /translation="MAGSGCAGAPPRFLAERGLMOVSRIGSSASVYRVCQG
 TPGSPGALAKOPLPGTTGAASAAEYGERKRALBQIGHRNIVLYGVTHFSP
 NPSRCLLELDVSESLVSSHQSGMMNIQHCADVDLALFLHBEVHADLK
 PNIIIMASNEKPKLIDPGLSPKGNQVXIQTGYRPAPEALONCIAQLQSDTE
 CTSADVLSIGIILIEFSGMKLKHVTSQEKANSASAIIDHIFASKAVVNAIDAYH
 LBDLAKSMHDEPSRIPAEMLCSPPFSIPAPRIEDLVMLPTVLRLLVLDYDL
 ENDEVEDVEDKEKCKYGPVSLVLPKRNPGQGVFEVANNAGDSAAKLTLGR
 MFDGFVAATFPFLSAIKRGLYQTL"
 ORIGIN
 Query Match 87.8%; Score 1106.4; DB 10; Length 3244;
 Best local Similarity 92.4%; Pred. No. 2,1e-272;
 Matches 1164; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
 QY 1 ATGGCGGAGTCCGGCTGCGCTGGGGCGGAGCCGCGCTTTCTGAGAGCTTCGGG 60
 Db 162 ATGGCGGAGTCCGGCTGCGCTGGGGCGGAGCCGCGCTTCCTGAGAGCTTCGGG 221
 QY 61 CGGCTGTGGCAGGTACAGAGCCGCTCTGGGTACGGCTCTCCGCTCGGTATAGGGTT 120
 Db 222 CGGCTGTGGCAGGTACAGAGCCGCTGGGTACGGCTCTCCGCTCGGTATAGGGTT 281
 QY 121 CGCTGTGGCAGCAACCTTGCTGCGCCCGCGGCGCTTCAGAGCAATTCCTGCGCAGGA 180
 Db 282 CGCTGTGGCAGCACTCAGAGCTCGCCCGCGGCGCTTCAGAGCAATTCCTGCGCAGGA 341
 QY 181 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGGTTTCCGAAAGAGAGGGCGGCTG 240
 Db 342 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGGTTTCCGAAAGAGAGGGCGGCTG 401
 QY 241 GAAACGTTGACAGGTACAGAAACATCGTACTTTGTATGAGTGTATCAATCACTTT 300
 Db 402 GAGCAGTTGACAGGTACAGAAACATCGTACTTTGTATGAGTGTATCAATCACTTT 461
 QY 301 TCTCCAAATGTCATCAACGCTGTCTGTTGCTTGAACCTCGATGTGAGTGTTCGGA 360
 Db 462 TCTCCAAATGTCATCAACGCTGTCTGTTGCTTGAACCTCGATGTGAGTGTTCGGA 521
 QY 361 TTGCTTTATATTCAGTACACAGAGGTGTTCATGTGATGATACAGATTCGCGCGA 420
 Db 522 TTGCTTTATATTCAGTACACAGAGGTGTTCATGTGATGATACAGATTCGCGCGA 581
 QY 421 GATGTTTGAAGCCCTGCTTTCTTCAATCATGAGGCTATGTCCATGCGGACTCAAA 480
 Db 582 GATGTTTGAAGCCCTGCTTTCTTCAATCATGAGGCTATGTCCATGCGGACTCAAG 641
 QY 481 CCACGTAAATATGTTGAGTGCAGAGATGAATGTTTAACTGATGACTTTGGAATT 540
 Db 642 CCACGTAAATATGTTGAGTGCAGAGATGAATGTTTAACTGATGACTTTGGAATT 701
 QY 541 ACCTTCAAGAGGACATCAGAGTAAATATATTCAGACAGAGGATTCGGCTCCA 600
 Db 702 ACCTTCAAGAGGACATCAGAGTAAATATATTCAGACAGAGGATTCGGCTCCA 761
 QY 601 GAAGCAAAATTCGAAATTTGCTTGGCCAGAGCTGCGCTGACAGTATCAGAAATGTACC 660
 Db 762 GAAGCAAAATTCGAAATTTGCTTGGCCAGAGCTGCGCTGACAGTATCAGAAATGTACC 821
 QY 661 TCAGCGTTGATCTGAGAGCTAGAGATCAATTTAATGGAATGTTTCAGGAATGAAA 720
 Db 822 TCAGCGTTGATCTGAGAGCTAGAGATCAATTTAATGGAATGTTTCAGGAATGAAA 881
 QY 721 CTGAAACATACAGTACATCTCAGAGATGAAAGGCAACAGTTCGTATTAATGATCAC 780
 Db 882 CTGAAACATACAGTACATCTCAGAGATGAAAGGCAACAGTTCGTATTAATGATCAT 941
 QY 781 AATTTTGCAGTAAAGCAGTGTGAATGCCGAATTCAGGCTTACCTTAAGAGACTT 840
 Db 942 AATTTTGCAGTAAAGCAGTGTGAATGCCGAATTCAGGCTTACCTTAAGAGACTT 1001

QY 841 ATCAAAAGCATGCTCATGATGATCAAGCAGAAATCTCTGAAATGGCATTGTGC 900
DB 1002 ATCAAAAGCATGCTCATGATGATCAAGCAGAAATCTCTGAAATGGCATTGTGC 1061
QY 901 AGCCCATTTCTTATGCAATCTCTTGGCCCTCATATTTGAAGATCTGCTATCTTCCACT 960
DB 1062 AGCCCATTTCTTATGCAATCTCTTGGCCCTCATATTTGAAGATCTGCTATCTTCCACT 1121
QY 961 CCAGTCTAAGATGCTGTAATGTCTGATGATGATATTTATTTGGAATGAAGAATAT 1020
DB 1122 CCAGTCTAAGATGCTGTAATGTCTGATGATGATATTTATTTGGAATGAAGAATAT 1181
QY 1021 GAAGATGTTGTAGAGATGTAAGAGAGAGATGTCAAAATATATGACCAAGTATCTTA 1080
DB 1182 GAAGATGTTGTAGAGATGTAAGAGAGAGATGTCAAAATATATGACCAAGTATCTTA 1241
QY 1081 CTGTGTTCCAAAGAAATCTCTGCAAGACAAAGTCTTTGTATGATGAAATGCTGTG 1140
DB 1242 CTGTGTTCCAAAGAAATCTCTGCAAGACAAAGTCTTTGTATGATGAAATGCTGTG 1301
QY 1141 GATTCCAAAGCTGCGAGAAATTAATGATGATGAAAGATGTTGATGGAAGTTGTTG 1200
DB 1302 GATTCCAAAGCTGCGAGAAATTAATGATGATGAAAGATGTTGATGGAAGTTGTTG 1361
QY 1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTA 1260
DB 1362 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACTTGTCTTA 1421

RESULT 13
MMPKIS 1254 bp mRNA linear ROD 08-JUL-2002
LOCUS Mus musculus partial mRNA for KIS protein kinase (Kis gene).
DEFINITION Y10725.2 GI:21726712
VERSION Y10725.2 GI:21726712
KEYWORDS Kis gene; protein kinase.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1254)
Maucuer, A., Ozon, S., Manceau, V., Gavet, O., Lawler, S., Curmi, P. and Sobel, A.
KIS is a protein kinase with an RNA recognition motif
J. Biol. Chem. 272 (37), 23151-23156 (1997)
MEDLINE 97435279
PUBMED 9287318

REFERENCE

REFERENCE
AUTHORS Maucuer, A.
JOURNAL Direct Submission
TITLES Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du
Fer Moulin, Paris, 75005, FRANCE
REMARK revised by (4)
3 (bases 1 to 1254)

REMARK
AUTHORS Maucuer, A.
JOURNAL Direct Submission
TITLES Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du
Fer Moulin, Paris, 75005, FRANCE
COMMENT On Jul 10, 2002 this sequence version replaced gi:1806129.
Overlaps with related sequences X82320, X10725.

FEATURES

source
1. 1254
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="K2"
/tissue_type="brain"
/clone_lib="lambda gt10"
/dev_stage="neonate"
1. 1254
/gene="Kis"
11. >1254
/gene="Kis"

/codon_start=1
/product="KIS protein kinase"
/protein_id="CAA71714.2"
/db_xref="GI:21726713"
/db_xref="GOA:P97343"
/db_xref="Swiss-Prot:P97343"
/translation="MAGSGCAGAGAPPEPFLAEFRLMVOYHLSGSSASVRYRCCG
TPGSPGALKQPLPGTTGAAASAEYFPERKRALAQHGRNIVLYGVFTTHFS
NPSRCLLELDIVSELLIYSSHQSSMMI QHCARDVLEALAFHHEGVADLK
PNILMSAENECFKLIDFGLSPENODVRYOTDGYAPAEALQCLAQGLQSDTE
CTSAVDIMSIGITILEMRSQMKLKHITVSQEKLSAIIIDHIFSKAVVNAIPAYH
LRDLKSMHEDHPKRIIPAEMLCSPPFSIPRAPIEDLVNIPPVATLAVDDYL
ENEDYEIVDEVEKKECCQYFPVSLVLPKENPBGQVFEYVNAAGDSKAOKLITGR
MEDGFVATFVPELPSAYRGLY"

ORIGIN

Query Match 87.2%; Score 1098.4; DB 10; Length 1254;
Best Local Similarity 92.7%; Pred. No. 1.9e-270;
Matches 1153; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 1 ATGCGCGGATCCGGCTGCGCTGCGGCGCGGAGCCGCGCTTTCTGAGGCGCTTCGGG 60
DB 11 ATGCGCGGATCCGGCTGCGCTGCGGCGCGGAGCCGCGCTTCCTGAGGCGCTTCGGG 70
QY 61 CGGCTGTGCGAGGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTCGGTGTATCGGCTT 120
DB 71 CGGCTGTGCGAGGTACAGAGCCGTCTGGGTAGCGGCTCTCGGCTCGGTGTATCGGCTG 130
QY 121 CGGCTGTGCGAGCCCTGCGCTGCGGCGCGGCGCGCTTCAGGCGAGTCTTGCGCGCAGGA 180
DB 131 CGGCTGTGCGGTACCCGGGCTGCGGCGCGGCGCGCTTCAGGCGAGTCTTGCGCTCGGGA 190
QY 181 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 240
DB 191 ACCACCGGGGCTGCGGCTCTGCGCGCGAGTATGTTTCCGAAAGAGAGGCGGCGCTG 250
QY 241 GAACAGTTGCGAGGTACAGAAACATCGTACTTGTATGAGGTGTTCAATCCACTTT 300
DB 251 GAGCAGTTGCGAGGTACAGAAACATCGTACTTGTATGAGGTGTTCAATCCACTTT 310
QY 301 TCTCCAAATGCGCATCAGCGTGTCTGTTGCTTGAACCTCGTAGATGCTGTTTCGGA 360
DB 311 TCTCCAAATGCGCATCAGCGTGTCTGTTGCTTGAACCTCGTAGATGCTGTTTCGGA 370
QY 361 TTGCTCTTATTTCCAGTACCAAGGTTGTTCCATGTGATGATACAGATTGCGGCCGA 420
DB 371 TTGCTCTTATTTCCAGTACCAAGGTTGTTCCATGTGATGATACAGATTGCGGCCGA 430
QY 421 GATGTTTGAAGCCCTTCTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 480
DB 431 GATGTTTGAAGCCCTTCTTCTTCATCATGAGGCTATGTCATGCGGACTCAAA 490
QY 481 CCAGCTTAATATGTTGAGTGCAGAGAAATGTTTAACTATGACTTGGAGCT 540
DB 491 CCAGCTTAATATGTTGAGTGCAGAGAAATGTTTAACTATGACTTGGAGCT 550
QY 541 AGCTTCAAGAGGCAATCAGATGTAAGTATTTACAGACAGAGGATGCGGCTCA 600
DB 551 AGCTTCAAGAGGCAATCAGATGTAAGTATTTACAGACAGAGGATGCGGCTCA 610
QY 601 GAAGCAGATTGCAAAATGCTTGGCCGAGGCTGCGTGCAGAGTATACAGATGATCC 660
DB 611 GAAGCAGATTGCAAAATGCTTGGCCGAGGCTGCGTGCAGAGTATACAGATGATCC 670
QY 661 TCAAGCTGTGATCTGTGAGCCTTAGAATCATTTTACTGGAATGTTCTCAGGAATGAAG 720
DB 671 TCAAGCTGTGATCTGTGAGCCTTAGAATCATTTTACTGGAATGTTCTCAGGAATGAAG 730
QY 721 CTGAAGCATACAGTCAAGATCTCAGAGATGAAGGCAACAGTTCTGCTATTAATGATCAC 780
DB 731 CTGAAGCATACAGTCAAGATCTCAGAGATGAAGGCAACAGTTCTGCTATTAATGATCAT 790
QY 781 AATATTGCCAGTAAAGCAGTGTGTAATGCCGCAATTCAGCTATCACTTAAGAGACTT 840

```

Db      791 ATATTGGCAAGTAAGCAATGGATGCGCAATTCAGCCATACCTCAGAGACCTT 850
Qy      841 ATCAAAAGCATGCTTCATGATGATCCAGCAGAAATTCCTGCTGAATGGCATTTGTC 900
Db      851 ATCAAAAGCATGCTTCATGATGATCCAGCAGAAATTCCTGCTGAATGGCATTTGTC 910
Qy      901 AGCCCATTTTGAAGCATTCCTTTGGCCCTCAATTTGAAGATGTCGTCATGCTCCACT 960
Db      911 AGCCCATTTTGAAGCATTCCTTTGGCCCTCAATTTGAAGATGTCGTCATGCTCCACT 970
Qy      961 CCAGTCTAAGCATGCTGCAATGCTGATGATGATATTTCTGGGATGAGAGAAATAT 1020
Db      971 CCAGTCTAAGCATGCTGCAATGCTGATGATGATATTTCTGGGATGAGAGAAATAT 1030
Qy      1021 GAAGATGTTGAGAGATGTAAGAGAGAGTGTCAAAAATATGACCACTGATCTCTA 1080
Db      1031 GAAGATGTTGAGAGATGTAAGAGAGAGTGTCAAAAATATGACCACTGATCTCTG 1090
Qy      1081 CTGTGTCGAAGAAATCTGCGACAGAGCAAGTCTTTGATGATGAAATGCTGGT 1140
Db      1091 CTGTGTCGAAGAAATCTGCGACAGAGCAAGTCTTTGATGATGAAATGCTGGT 1150
Qy      1141 GATTCCAAAGCTGCGGAGAAATTAATGATGATGAGATGTTGATGAGAGATTTGTTG 1200
Db      1151 GATTCCAAAGCTGCTGAGAAATGCTGATGATGAGAGATGTTGATGAGAGATTTGTTG 1210
Qy      1201 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTGTA 1244
Db      1211 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTTXX 1254

```

```

RESULT 14
BC026046      1367 bp      mRNA      linear      PRI 30-JUN-2004
LOCUS      Homo sapiens kinase interacting with leukemia-associated gene
DEFINITION      (lethalmn), mRNA (cDNA clone MGC:39391 IMAGE:4649261), complete
cds.

```

```

ACCESSION      BC026046      GI:19684094
VERSION      BC026046.1
KEYWORDS      MGC.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

1 (bases 1 to 1367)
Straussberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, K.H., Schaefer, C.F., Bhat, N.K.,
Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Sheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mulhaly, S.J., Bosak, S.A., McKean, P.J.,
McKernan, K.J., Malek, J.A., Gamaralle, P.H., Richards, S.,
Mortley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, X., Gibbs, R.A.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Bouffard, G.G., Blakeley, R.W., Madan, A., Young, A.C., Shevchenko, Y.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smallus, D.E.,
Schnurch, A., Schein, J.B., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

```

```

JOURNAL      PUBMED
PUBMED      12477932
REFERENCE      2 (bases 1 to 1367)
AUTHORS      Straussberg, R.
TITLE      Direct Submission
JOURNAL      Submitted (20-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

```

```

REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Telka Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu,
Parvaneh Saedi, Ur Santos, Angelique Schmech, Ursula Skalske,
Duane Smallus, Jeff Stott, Miranda Tsai, George Yang, Jacques
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

```

```

FEATURES
source
1..1367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:39391 IMAGE:4649261"
/tissue_type="Eye, retinoblastoma"
/clone_id="NM_119181"
/lab_host="DH10B-R_16"
/note="vector: pOTB7"
1..1367
/gene="KIS"
/db_xref="LOCUSID:127933"
159..1193
/gene="KIS"
/codon_start=1
/product="kinase interacting seahmin"
/protein_id="AAH26046.1"
/db_xref="GI:19684095"
/db_xref="LOCUSID:127933"
/translaton="MAGSCAMGAPRFLAFLGMLYQSLGSSASVRYRCC
NGSPGALKQFLPGTTGAASAAYEPRERKALBOLQGRNVTLLGVTFHSP
NPSPCLLELDVSVBLLYSHQGSMWMTQICARDVLAFLHREGVADLK
PNPILMSABNECFKLIDFGLSPKSNQOVKTYOTGYPAPAEALQNLQSDTE
CSAVDMSIGITILEMFSGMKLTHTVSQEKANSATIDHIFASKVNVNAIPAYH
LRDLKSMHDPSPRIRAEALCSPSPISIPAPAHIEDLVMLPTVLAUNLVDDYL
ENEREYEGIC"

```

```

gene
cds

```

```

ORIGIN

```

```

Query Match      84.6%; Score 1065.6; DB 9; Length 1367;
Best Local Similarity 92.6%; Pred. No. 5; le-262;
Matches 1167; Conservative 0; Mismatches 4; Indels 89; Gaps 1;
Qy      1 ATGGCGGAGTCCGCGCTGCGCTGCGGCGCGGAGCGCGCGCTTTCTGAGAGCGCTTGGG 60
Db      159 ATGGCGGAGTCCGCGCTGCGCTGCGGCGCGGAGCGCGCGCTTTCTGAGAGCGCTTGGG 218
Qy      61 CGGCTGTGCGAGTACAGAGCGCGCTGCGGTAACGGCTCTCCGCGCTCGGTATTCGGCTT 120
Db      219 CGGCTGTGCGAGTACAGAGCGCGCTGCGGTAACGGCTCTCCGCGCTCGGTATTCGGCTT 278
Qy      121 CGCTGTGCGGGAACCGTGGCTCGCCCGCGCGCGCTCAGGCACTTCTTGGCGCAGGA 180
Db      279 CGCTGTGCGGGAACCGTGGCTCGCCCGCGCGCGCTCAGGCACTTCTTGGCGCAGGA 338
Qy      181 ACCACCGGCGCTGCGGCTCTGCGCGCGGAGTATGCTTTCCGAAAGAGAGGCGCGCTG 240
Db      339 ACCACCGGCGCTGCGGCTCTGCGCGCGGAGTATGCTTTCCGAAAGAGAGGCGCGCTG 398

```


Db 423 AACTATATTCAGACAGACGGGTATCGGGCTCCAGAAACAGAAATTGCAAAATTGCTTGCC 482
QY 628 CAGGCTGGCCTGCAGATGATACAGAAATGATACCTCAGCTGTGATCTGTGAGCCCTAGGA 687
Db 483 CAGGCTGGCCTGCAGATGATACAGAAATGATACCTCAGCTGTGATCTGTGAGCCCTAGGA 542
QY 688 ATCATTTTACTGGAATGTTCTCAGAAATGAACTGAAACATACATCAGATCTCAGAA 747
Db 543 ATCATTTTACTGGAATGTTCTCAGAAATGAACTGAAACATACATCAGATCTCAGAA 602
QY 748 TGGAAAGCAAAACAGTTCTGCTATTATTGATCAGATTTGCGAATAAGCAGTGTGAAT 807
Db 603 TGGAAAGCAAAACAGTTCTGCTATTATTGATCAGATTTGCGAATAAGCAGTGTGAAT 662
QY 808 GCCGCAATTCACACCTATCAGCTAAGAGACCTTATCAAAGCATGCTTCAATGATCCA 867
Db 663 GCCGCAATTCACACCTATCAGCTAAGAGACCTTATCAAAGCATGCTTCAATGATCCA 722
QY 868 AGCAGAGAATTCCTGCTGAATGCAATGTCAGCCCATTTTAGCATTCCTTTGCG 927
Db 723 AGCAGAGAATTCCTGCTGAATGCAATGTCAGCCCATTTTAGCATTCCTTTGCG 782
QY 928 CCTCATATTGAAGATCTGTCATGCTTCCCATCCAGTGTAAAGACTGCTGAATGTCTG 987
Db 783 CCTCATATTGAAGATCTGTCATGCTTCCCATCCAGTGTAAAGACTGCTGAATGTCTG 842
QY 988 GATGATGATTAATCTTGGAATGAAAGAAATATGAAGATGTTGTAAGATGTAAGAG 1047
Db 843 GATGATGATTAATCTTGGAATGAAAGAAATATGAAGATGTTGTAAGATGTAAGAG 902
QY 1048 GAGTGTCAAAAATATGGAACAGATGATCTCTACTTGTTCAAAAGAAAATCTGGCAGA 1107
Db 903 GAGTGTCAAAAATATGGAACAGATGATCTCTACTTGTTCAAAAGAAAATCTGGCAGA 962
QY 1108 GGAACAAGCTTTGTTGATGATGCAAAATGCTGTGATTCAAAAGCTGGCAGAAATTACTG 1167
Db 963 GGAACAAGCTTTGTTGATGATGCAAAATGCTGTGATTCAAAAGCTGGCAGAAATTACTG 1022
QY 1168 ACTGGAAGGATGTTTGAATGGGAAGTTTGTGTGCTACATTCACCCGCTGAGTGCCTAC 1227
Db 1023 ACTGGAAGGATGTTTGAATGGGAAGTTTGTGTGCTACATTCACCCGCTGAGTGCCTAC 1082
QY 1228 AAGAGGGGATATCTGTATCAAACTTGCTTTAA 1260
Db 1083 AAGAGGGGATATCTGTATCAAACTTGCTTTAA 1115

Search completed: November 29, 2004, 14:49:20
Job time : 5514 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 11:23:25 ; Search time 653.5 Seconds
(without alignments)
10121.301 Million cell updates/sec

Title: US-10-798-532-3

Perfect score: 1260
Sequence: 1 atggggggatccgctgcgc.....tgatcaacctgtcttaa 1260

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|--------|-------------|-----------|----|--------------------|
| 1 | 1260 | 100.0 | 1260 | 3 | AAZ51356 Human tta |
| 2 | 1258.4 | 99.9 | 1260 | 3 | AAZ51355 Human wil |
| 3 | 1256.8 | 99.7 | 2008 | 6 | ABQ93183 Human cod |
| 4 | 1255.2 | 99.6 | 1260 | 4 | AAI66828 Human pro |
| 5 | 1255.2 | 99.6 | 2622 | 4 | AAI66828 Human pro |
| 6 | 1255.2 | 99.6 | 4065 | 5 | ADL45238 Human ova |
| 7 | 1255.2 | 99.6 | 4065 | 5 | ADL45238 Human ova |
| 8 | 1253.6 | 99.5 | 1260 | 4 | AAZ51356 Human tta |
| 9 | 1253.6 | 99.5 | 1260 | 4 | AAZ51356 Human tta |
| 10 | 1173.4 | 93.1 | 3998 | 12 | ADQ24401 Human sof |
| 11 | 1106.4 | 87.8 | 3244 | 10 | ADBS2857 Primary r |
| 12 | 458.6 | 36.4 | 529 | 4 | AAH70764 Human cer |
| 13 | 445.8 | 35.4 | 722 | 3 | AAA02535 Human col |
| 14 | 426.6 | 33.9 | 470 | 4 | AAH71966 Human cer |
| 15 | 375.2 | 29.8 | 461 | 4 | AAH69101 Human cer |
| 16 | 272.2 | 21.6 | 411 | 6 | ABQ99225 Human ORF |
| 17 | 203.8 | 16.2 | 231 | 3 | AAA01259 Human ORF |
| 18 | 177.4 | 14.1 | 419 | 5 | ADL73740 Human ova |
| 19 | 177.4 | 14.1 | 419 | 5 | ADL73740 Human ova |
| 20 | 176.2 | 14.0 | 546 | 6 | ABQ53127 Oligonuc1 |
| 21 | 176.2 | 14.0 | 546 | 6 | ABQ53126 Oligonuc1 |

| | | | | | | | |
|---|----|-------|------|------|----|----------|--------------------|
| C | 22 | 176 | 14.0 | 547 | 6 | ABQ44465 | Abq44465 Oligonuc1 |
| | 23 | 176 | 14.0 | 547 | 6 | ABQ44464 | Abq44464 Oligonuc1 |
| | 24 | 157.2 | 12.5 | 546 | 6 | ABQ53129 | Abq53129 Oligonuc1 |
| C | 25 | 157.2 | 12.5 | 546 | 6 | ABQ53128 | Abq53128 Oligonuc1 |
| | 26 | 157.2 | 12.5 | 547 | 6 | ABQ44463 | Abq44463 Oligonuc1 |
| C | 27 | 157.2 | 12.5 | 547 | 6 | ABQ44462 | Abq44462 Oligonuc1 |
| | 28 | 148.4 | 11.8 | 442 | 9 | ACH28557 | ACH28557 Human adu |
| | 29 | 85 | 6.7 | 198 | 5 | ADJ38977 | ADJ38977 Human ova |
| | 30 | 63.4 | 5.0 | 65 | 6 | ABN52650 | ABN52650 Mouse sp1 |
| | 31 | 63.2 | 5.0 | 549 | 10 | ADK58517 | ADK58517 Plant DNA |
| | 32 | 57.8 | 4.6 | 1196 | 12 | ADJ39614 | ADJ39614 Plant CDN |
| | 33 | 53.8 | 4.3 | 1364 | 12 | ADJ42189 | ADJ42189 Plant CDN |
| | 34 | 47.4 | 3.8 | 1747 | 6 | AAZ47544 | AAZ47544 Human DNA |
| | 35 | 47.4 | 3.8 | 1863 | 3 | AAZ47544 | AAZ47544 Human pro |
| | 36 | 47.4 | 3.8 | 2044 | 8 | ABX70932 | ABX70932 Novel hum |
| | 37 | 47.4 | 3.8 | 2887 | 12 | ADJ95566 | ADJ95566 Human Dyr |
| | 38 | 46.8 | 3.7 | 7565 | 12 | ADK70703 | ADK70703 Murine CR |
| | 39 | 45.4 | 3.6 | 1581 | 6 | ABK88905 | ABK88905 Murine md |
| | 40 | 44.4 | 3.5 | 2658 | 3 | AAA70213 | AAA70213 Plasmodiu |
| | 41 | 43.8 | 3.5 | 1584 | 6 | ABL58336 | ABL58336 Rat YAK1 |
| | 42 | 43 | 3.4 | 681 | 10 | ADK52877 | ADK52877 Plant DNA |
| C | 43 | 42.8 | 3.4 | 600 | 12 | ADJ44008 | ADJ44008 Plant CDN |
| | 44 | 42.4 | 3.4 | 1497 | 3 | AAA70154 | AAA70154 Plasmodiu |
| | 45 | 42.2 | 3.3 | 2095 | 3 | AAZ36941 | AAZ36941 Arabidops |

ALIGNMENTS

| | | |
|----------|--|---|
| RESULT 1 | AAZ51356 | AAZ51356 standard; DNA; 1260 BP. |
| ID | AAZ51356 | |
| XX | AAZ51356 | |
| AC | AAZ51356 | |
| XX | 06-JUN-2000 | (first entry) |
| DT | 06-JUN-2000 | (first entry) |
| XX | Human transdominant mutant serine/threonine kinase KIS (hKIS) gene. | |
| DE | KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase; | |
| XX | transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27; | |
| KW | modulator; treatment; cell proliferative disease; vascular disorder; | |
| KM | gene therapy; atherosclerosis; restenosis; ds. | |
| XX | Homo sapiens. | |
| OS | Homo sapiens. | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | CDS | 1..1260 |
| FT | | /*tag= a |
| FT | | /product= "Human transdominant mutant KIS protein" |
| FT | | /transl_except= (pos:556..561, Asn) |
| FT | misc_feature | 161 |
| FT | | /*tag= b |
| FT | | /note= "A to G substitution resulting in a Lys to Arg |
| FT | | change in the amino acid sequence" |
| XX | WO20001165-A1. | |
| XX | 02-MAR-2000. | |
| PD | 20-AUG-1999; | 99WO-US018903. |
| PF | 21-AUG-1998; | 98US-0097710P. |
| PR | | |
| XX | (NABE/) NABEL G J. | |
| PA | (NABE/) NABEL B G. | |
| XX | | |
| PI | Nabel GJ, Nabel EG; | |
| XX | | |
| XX | WPI; 2000-237648/20. | |
| DR | P-PDB; AAY70306. | |
| XX | | |
| PT | Novel serine/threonine kinase hKIS polynucleotides and polypeptides used | |

PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 PT proliferation.

PS Claim 7; Page 59; 70pp; English.

XX The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
 CC constructed by site directed mutagenesis. A single nucleotide
 CC substitution (A to G) results in a lysine to arginine change in the
 CC protein sequence. hKIS is a serine/threonine kinase, that acts as an
 CC inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
 CC controls cell proliferation and is localised predominantly in the
 CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
 CC G1 phase. The hKIS sequences are used to modulate cell proliferation and
 CC treat cell proliferative and vascular diseases. The polynucleotide
 CC sequence may be used in gene therapy to treat vascular disorders such as
 CC restenosis or atherosclerosis

SQ Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

Query Match 100.0%; Score 1260; DB 3; Length 1260;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGATCCGGCTGGCGCGCGGAGCCCGCGCTTTCTGAGAGCGCTTCGGG 60
 DB 1 ATGGCGGATCCGGCTGGCGCGCGGAGCCCGCGCTTTCTGAGAGCGCTTCGGG 60
 QY CGGCTGTGAGGATGAGAGCGCGCTGTGGAGCGCGCTTCGCGCTGTGATTCGGGTT 120
 DB CGGCTGTGAGGATGAGAGCGCGCTGTGGAGCGCGCTTCGCGCTGTGATTCGGGTT 120
 QY 121 CGCTGTGCGGCAACCTGTGCTGCGCGCGCGCGCTTCGAGGAGTTCTTCCGCGAGGA 180
 DB 121 CGCTGTGCGGCAACCTGTGCTGCGCGCGCGCGCTTCGAGGAGTTCTTCCGCGAGGA 180
 QY 181 ACCACCGGCGCTGCGCGCTTCGCGCGAGTATGTTCCGCAAGAGGCGCGCGCTG 240
 DB 181 ACCACCGGCGCTGCGCGCTTCGCGCGAGTATGTTCCGCAAGAGGCGCGCGCTG 240
 QY 241 GAACAGTTCGAGGATGAGAGCAATCGTATCTTGTATGAGAGTGTTCATCACTCTT 300
 DB 241 GAACAGTTCGAGGATGAGAGCAATCGTATCTTGTATGAGAGTGTTCATCACTCTT 300
 QY 301 TCTCCAAATGTGCGATCAGCTGTGTGTGCTTGAACCTCTGATGTCACTGTTTGGAA 360
 DB 301 TCTCCAAATGTGCGATCAGCTGTGTGTGCTTGAACCTCTGATGTCACTGTTTGGAA 360
 QY 361 TTGCTCTTATATCCAGTCAACGAGGTGTTCATGTGATGATGATGAGATTCGCGCGA 420
 DB 361 TTGCTCTTATATCCAGTCAACGAGGTGTTCATGTGATGATGATGAGATTCGCGCGA 420
 QY 421 GATGTTTGGAGGCGCTTGTCTTTCATCATGAGGCTATGTCATGCGAGCTCAAA 480
 DB 421 GATGTTTGGAGGCGCTTGTCTTTCATCATGAGGCTATGTCATGCGAGCTCAAA 480
 QY 481 CCAAGTATATTTGAGAGTCAAGATGAAATGTTTAACTCACTTGAAGCTT 540
 DB 481 CCAAGTATATTTGAGAGTCAAGATGAAATGTTTAACTCACTTGAAGCTT 540
 QY 541 AGCTTAAAGAGGAGATGAGATGATTAATTCAGACAGCGGATGCGGCTCCA 600
 DB 541 AGCTTAAAGAGGAGATGAGATGATTAATTCAGACAGCGGATGCGGCTCCA 600
 QY 601 GAAGCAGAAATGCAAAATGCTTGGCCAGAGCTGCGCTGAGATGATGAGATGATGAC 660
 DB 601 GAAGCAGAAATGCAAAATGCTTGGCCAGAGCTGCGCTGAGATGATGAGATGATGAC 660
 QY 661 TCAGCTGTGATGTGAGAGCTTGAAGATCAATTTTCTGAGAAATGTTTCAAGAAATGAAA 720
 DB 661 TCAGCTGTGATGTGAGAGCTTGAAGATCAATTTTCTGAGAAATGTTTCAAGAAATGAAA 720
 QY 721 CTGAAACATACATGATCTCAGAAATGAGCAACGTTCTGCTATATATGATCAC 780
 DB 721 CTGAAACATACATGATCTCAGAAATGAGCAACGTTCTGCTATATATGATCAC 780

DB 721 CTGAAACATACATGATCTCAGAAATGAGCAACGTTCTGCTATATATGATCAC 780
 QY 781 ATATTGGCCAGTAAAGCAGTGTGATGATGCGCAATTCACGCTTACCTAGAGACCTT 840
 DB 781 ATATTGGCCAGTAAAGCAGTGTGATGATGCGCAATTCACGCTTACCTAGAGACCTT 840
 QY 841 ATCAAAGCAGTCTTCATGATATATCCAGAGAGAAATTCCTGCTGAAATGCAATTCGTC 900
 DB 841 ATCAAAGCAGTCTTCATGATATATCCAGAGAGAAATTCCTGCTGAAATGCAATTCGTC 900
 QY 901 ACCCATTTTGAAGATTCCTTTGCCCCATATTTGAAGATCTGTCATGCTTCCACT 960
 DB 901 ACCCATTTTGAAGATTCCTTTGCCCCATATTTGAAGATCTGTCATGCTTCCACT 960
 QY 961 CCAAGTCTAAGATCTGTAATGTGTAATGATATCTTGGAAATGAGAGAAATAT 1020
 DB 961 CCAAGTCTAAGATCTGTAATGTGTAATGATATCTTGGAAATGAGAGAAATAT 1020
 QY 1021 GAAGATGTTTGAAGATGTAAGAGAGTCAAAATATGACCACTGCTATCTTA 1080
 DB 1021 GAAGATGTTTGAAGATGTAAGAGAGTCAAAATATGACCACTGCTATCTTA 1080
 QY 1081 CTGTTCCAAAGAAATCTGCGAGAGCAAGTCTTTGTGATGATGCAATGCTGCT 1140
 DB 1081 CTGTTCCAAAGAAATCTGCGAGAGCAAGTCTTTGTGATGATGCAATGCTGCT 1140
 QY 1141 GATTCMAAGTCTGCGAGAAATTAAGTCTGAGAGATGTTTGAAGAGATTGTTGTG 1200
 DB 1141 GATTCMAAGTCTGCGAGAAATTAAGTCTGAGAGATGTTTGAAGAGATTGTTGTG 1200
 QY 1201 GGTACATTTTACCGCTGAGTCCCTCAAGAGGAGATATCTGATCAAACTTGTGCTTAA 1260
 DB 1201 GGTACATTTTACCGCTGAGTCCCTCAAGAGGAGATATCTGATCAAACTTGTGCTTAA 1260

RESULT 2
 AA251355
 ID AA251355 standard; DNA; 1260 BP.
 AC AA251355;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Human wild type serine/threonine kinase KIS (hKIS) gene.
 XX
 KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
 KW dominant gene; cyclin-dependent kinase inhibitor; CKI; p27; modulator;
 KW treatment; cell proliferative disease; vascular disorder; gene therapy;
 KW atherosclerosis; restenosis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /tag= a
 FT /product= "Human wild type KIS protein"
 FT /transl_except= (pos:556..561, Asn)
 XX
 PN WO20001165-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 20-AUG-1999; 99WO-US018903.
 XX
 PR 21-AUG-1998; 98US-0097710P.
 XX
 PA (NABE/) NABEL G J.
 XX (NABE/) NABEL E G.
 XX
 PI Nabel GJ, Nabel EG;
 XX
 DR WPI: 2000-237648/20.
 DR P-PDB; AAT70305.

XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.

XX Claim 1; Page 56; 70pp; English.

CC The present DNA sequence encodes a wild type human KIS (hKIS), isolated
CC from a human B-cell library, using a yeast two hybrid screening system.
CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
CC that controls cell proliferation and is localised predominantly in the
CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
CC G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
CC The hKIS sequences are used to modulate cell proliferation and treat cell
CC proliferative and vascular diseases. The polynucleotide sequence may be
CC used in gene therapy to treat vascular disorders such as restenosis or
CC atherosclerosis

XX Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;

Query Match 99.9%; Score 1258.4; DB 3; Length 1260;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCGCGGATCCGCGCTGCGCGCGCGGAGCGCGCGGCTTTCTGAGGCGCTTCGCGG 60
DB 1 ATGCGCGGATCCGCGCTGCGCGCGCGGAGCGCGCGGCTTTCTGAGGCGCTTCGCGG 60
QY 61 CGGCTGCGCAGGTACAGAGCCGCTGCGGTAGCGGCTCCTCGCGCTCGGTATCCGGTT 120
DB 61 CGGCTGCGCAGGTACAGAGCCGCTGCGGTAGCGGCTCCTCGCGCTCGGTATCCGGTT 120
QY 121 CGCTGCTGCGGCAACCTGCTGCGCGCGCGCGCGCTTCAAGGCACTTTCGCGCAGGA 180
DB 121 CGCTGCTGCGGCAACCTGCTGCGCGCGCGCGCGCTTCAAGGCACTTTCGCGCAGGA 180
QY 181 ACCACCGGGGCTGCGCGCTGCGCGCGAGTATGTTTCCGCAAGAGAGGCGCGCTG 240
DB 181 ACCACCGGGGCTGCGCGCTGCGCGCGAGTATGTTTCCGCAAGAGAGGCGCGCTG 240
QY 241 GAACAGTTGCGAGGTCACAGAAACATCGTACTTGTATGAGTGTATCAATCCACTTT 300
DB 241 GAACAGTTGCGAGGTCACAGAAACATCGTACTTGTATGAGTGTATCAATCCACTTT 300
QY 301 TCTCCAAATGTGCAATCAGCTGTCTGTGCTTGAACCTCTGATGTCAGTGTTCGGA 360
DB 301 TCTCCAAATGTGCAATCAGCTGTCTGTGCTTGAACCTCTGATGTCAGTGTTCGGA 360
QY 361 TTGCTCTTATATTCACAGTCAACGAGGTGTTCATGTGATGATATACAGATTGCGCCGA 420
DB 361 TTGCTCTTATATTCACAGTCAACGAGGTGTTCATGTGATGATATACAGATTGCGCCGA 420
QY 421 GATGTTTGGAGGCGCTGCTTTTCTTCATCATGAGGCGATATGTCATGCGCACTCAA 480
DB 421 GATGTTTGGAGGCGCTGCTTTTCTTCATCATGAGGCGATATGTCATGCGCACTCAA 480
QY 481 CCAAGTAACTATTTGAGTGCAGAGATGAATGTTTAACTCATTTGAATT 540
DB 481 CCAAGTAACTATTTGAGTGCAGAGATGAATGTTTAACTCATTTGAATT 540
QY 541 AGCTTCAAAAGAGCATCGAGTATTAATTAATTCAGACAGCGGTATCGGGCTCCA 600
DB 541 AGCTTCAAAAGAGCATCGAGTATTAATTAATTCAGACAGCGGTATCGGGCTCCA 600
QY 601 GAAGCAGAAATTTGCTTGGCCAGAGCTGCGCTGAGAGTATACAGATGTACC 660
DB 601 GAAGCAGAAATTTGCTTGGCCAGAGCTGCGCTGAGAGTATACAGATGTACC 660
QY 661 TCAAGCTGTGATCTGTGAGCCTTGAAGATCAATTTTACGAAATGTTCTCAGAAATGAAA 720
DB 661 TCAAGCTGTGATCTGTGAGCCTTGAAGATCAATTTTACGAAATGTTCTCAGAAATGAAA 720

QY 721 CTGAATACATACATGATCTCAGGAATGAAAGCAACAGTTCTCTATTTATGATCAG 780
DB 721 CTGAATACATACATGATCTCAGGAATGAAAGCAACAGTTCTCTATTTATGATCAG 780
QY 781 ATATTGGCAGTAAAGCAGTGTGAAATCCGCAATTCAGGCTATCACTTAAGACCTT 840
DB 781 ATATTGGCAGTAAAGCAGTGTGAAATCCGCAATTCAGGCTATCACTTAAGACCTT 840
QY 841 ATCAAAAGCATGCTTCATGATGATCCAGAGAAATTCCTGCTGAATGAGCATTTGTC 900
DB 841 ATCAAAAGCATGCTTCATGATGATCCAGAGAAATTCCTGCTGAATGAGCATTTGTC 900
QY 901 AGCCCATTTCTTGAAGATTCCTTTTCCCTCATATTTGAAGATCTGTCATGCTCCACT 960
DB 901 AGCCCATTTCTTGAAGATTCCTTTTCCCTCATATTTGAAGATCTGTCATGCTCCACT 960
QY 961 CCAAGTCTAAGACTGCTGATATGCTGATGATGATTTATCTTGGAAATGAAGAAATAT 1020
DB 961 CCAAGTCTAAGACTGCTGATATGCTGATGATGATTTATCTTGGAAATGAAGAAATAT 1020
QY 1021 GAAAGTGTGTGAAGATGTAAGAGAGAGTGTCAAAAATATGACACAGTGTATCTCTA 1080
DB 1021 GAAAGTGTGTGAAGATGTAAGAGAGAGTGTCAAAAATATGACACAGTGTATCTCTA 1080
QY 1081 CTGTGTTCCAAAGAAATCTGCGCAGAGACAAAGTCTTGTGATGATGCAAAATGCTGCT 1140
DB 1081 CTGTGTTCCAAAGAAATCTGCGCAGAGACAAAGTCTTGTGATGATGCAAAATGCTGCT 1140
QY 1141 GATTCCAAAGCTGCGCAGAAATTAATCTGACTGGAAGATGTTTGAATGGAAGTTTGTG 1200
DB 1141 GATTCCAAAGCTGCGCAGAAATTAATCTGACTGGAAGATGTTTGAATGGAAGTTTGTG 1200
QY 1201 GCTACATCTACCGCTGAGTGTCTTCAAGAGGGGATATCTGTATCAAACTTGTCTTAA 1260
DB 1201 GCTACATCTACCGCTGAGTGTCTTCAAGAGGGGATATCTGTATCAAACTTGTCTTAA 1260

RESULT 3
ID ABQ9383 standard; cDNA; 2008 BP.
XX ABQ9383;
AC XX
XX XX
DT 25-FEB-2003 (first entry)
XX XX
XX Human coding sequence SEQ ID 116.
DE XX
XX Human; expressed sequence tag; EST; haematopoietic disorder;
KW central nervous system disease; viral infection;
KW peripheral nervous system disease; non-healing wound; infectious disease;
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
KW fungal infection; autoimmune disorder; coagulation disorder; noctropic;
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
KW cytotoxic; hemostatic; virucide; antibacterial; fungicide;
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.
XX
OS Homo sapiens.
PN WO200259260-A2.
XX
PD 01-AUG-2002.
XX
PF 16-NOV-2001; 2001WO-US042950.
XX
PR 17-NOV-2000; 2000US-00714936.
XX
PA (HYSE-) HYSO INC.
XX
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX MPI; 2002-590824/63.
DR N-PDB; ABP64797.

XX New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.

XX Claim 1, SEQ ID NO 116; 394pp; English.

XX The present invention relates to novel human coding sequences (AB093268-
 CC AB093608) and protein (ABP6482-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a locipotentia or pluripotentia state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPD at
 CC ftp.wipo.int/pub/published_pct_sequences

SO Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

Query Match 99.7%; Score 1256.8; DB 6; Length 2008;

Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1258; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCGGAGATCCGGCTGGGCGGCGGAGAGCGCGGTTTGTGAGGCGCTTCGGG 60
 DB 79 ATGGCGGAGATCCGGCTGGGCGGCGGAGAGCGCGGTTTGTGAGGCGCTTCGGG 138
 QY 61 CGGCTGTGGCAGGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGATGATCGGATT 120
 DB 139 CGGCTGTGGCAGGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGATGATCGGATT 198
 QY 121 CGGCTGTGGCAGGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGATGATCGGATT 180
 DB 199 CGGCTGTGGCAGGTACAGAGCCGCTGTGGTAGCGGCTCCGCTCGGATGATCGGATT 258
 QY 181 ACCACCGGGGCTGGGCGCTGTGGTAGCGGCTCCGCTCGGATGATCGGATT 240
 DB 259 ACCACCGGGGCTGGGCGCTGTGGTAGCGGCTCCGCTCGGATGATCGGATT 318
 QY 241 GAACAGTTCAGGGGTACAGAAACATGATCTGTATGAGAGTGTTCATCACTCTT 300
 DB 319 GAACAGTTCAGGGGTACAGAAACATGATCTGTATGAGAGTGTTCATCACTCTT 378
 QY 301 TCTCCAAATGTGCGATCAGCTGTGTGTGTTGAATCTCTGATGATCGAGTTCGGA 360
 DB 379 TCTCCAAATGTGCGATCAGCTGTGTGTGTTGAATCTCTGATGATCGAGTTCGGA 438
 QY 361 TTGCTCTTATATTCAGGTACAGAGGTTGTTCATGATGATGATGATGATGATGATG 420
 DB 439 TTGCTCTTATATTCAGGTACAGAGGTTGTTCATGATGATGATGATGATGATGATG 498
 QY 421 GATGTTTGGAGGCGCTGCTTTCTTCAATCATGAGGCTATGCGAGCTCAAA 480
 DB 499 GATGTTTGGAGGCGCTGCTTTCTTCAATCATGAGGCTATGCGAGCTCAAA 558
 QY 481 CCACTTAACATATTTGAGAGTGAGAGAAATGTTTAACTCATGATCTTGGACTT 540

DB 559 CCACGTAACTATTTGTGAGAGTGACAGAGATGAAATGTTTAACTCATGACTTGGACTT 618
 QY 541 AGCTTCAAGAAAGGAGATCAGAGATGTAAGTATTTACAGACAGCGGATCGGCTCA 600
 DB 619 AGCTTCAAGAAAGGAGATCAGAGATGTAAGTATTTACAGACAGCGGATCGGCTCA 678
 QY 601 GAAGCAGATTTGCAAAATTTGCTTGGCCAGGCTGGCTGAGAGTATACAGATGATAC 660
 DB 679 GAAGCAGATTTGCAAAATTTGCTTGGCCAGGCTGGCTGAGAGTATACAGATGATAC 738
 QY 661 TAGCTGTGATCTGTGAGAGCTTAGAGATCATTTTACAGAAATTTCTCAGAAATGAA 720
 DB 739 TAGCTGTGATCTGTGAGAGCTTAGAGATCATTTTACAGAAATTTCTCAGAAATGAA 798
 QY 721 CTGAAACATACAGTACAGATCTCAGAGATGAAAGGCAACAGTTCTGATTTATGATAC 780
 DB 799 CTGAAACATACAGTACAGATCTCAGAGATGAAAGGCAACAGTTCTGATTTATGATAC 858
 QY 781 ATATTTGCGAGTAAAGCAGTGTGATGCGCAATTCAGGCTATCAGCTTAAGAGACCTT 840
 DB 859 ATATTTGCGAGTAAAGCAGTGTGATGCGCAATTCAGGCTATCAGCTTAAGAGACCTT 918
 QY 841 ATCAAAAGCATGCTTCAATGATGATCCAGCAGAAATTTCTGCTGAATGCAATTTGTC 900
 DB 919 ATCAAAAGCATGCTTCAATGATGATCCAGCAGAAATTTCTGCTGAATGCAATTTGTC 978
 QY 901 ACCCATTTCTTAAAGATCTTTTCCCTCATATTTGAAGATCTGATGATGCTTCCACT 960
 DB 979 AGCCATTTCTTAAAGATCTTTTCCCTCATATTTGAAGATCTGATGATGCTTCCACT 1038
 QY 961 CCAAGTCTAAGACTGTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 1039 CCAAGTCTAAGACTGTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1098
 QY 1021 GAAGATGTTGTAAGATGTAAGAGAGTCTCAAAATATGACAGTGTATCTCTA 1080
 DB 1099 GAAGATGTTGTAAGATGTAAGAGAGTCTCAAAATATGACAGTGTATCTCTA 1158
 QY 1081 CTTGTTCCAAAGAAATCTGTGAGAGCAAGTCTTGTGATGATGATGATGATGATGAT 1140
 DB 1159 CTTGTTCCAAAGAAATCTGTGAGAGCAAGTCTTGTGATGATGATGATGATGATGAT 1218
 QY 1141 GATTCAAAGCTGCGCAGAAATTAAGTCTGAGAGATGTTGATGAGAGATGTTGTTG 1200
 DB 1219 GATTCAAAGCTGCGCAGAAATTAAGTCTGAGAGATGTTGATGAGAGATGTTGTTG 1278
 QY 1201 GCTACATTTACCGCTGAGTCCCTAACAAGAGGATATCTGTATCAAACTTGTCTTA 1260
 DB 1279 GCTACATTTACCGCTGAGTCCCTAACAAGAGGATATCTGTATCAAACTTGTCTTA 1338

RESULT 4
 AA166829
 ID AA166829 standard; cDNA; 1260 BP.
 AC AA166829;
 XX
 DT 07-JAN-2002 (first entry)
 XX
 DE Human protein kinase polypeptide 13887 coding sequence.
 XX
 KW Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
 KW cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatotropic;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "protein kinase 13887"
 XX

| | |
|----|---|
| Pt | MO200173050-AZ. |
| XX | |
| PD | 04-OCT-2001. |
| XX | |
| PF | 23-MAR-2001; 2001WO-US009483. |
| PR | 24-MAR-2000; 2000US-0191846P. |
| XX | |
| PA | (MILL-) MILLENNIUM PHARM INC. |
| XX | |
| P1 | Meyers R; |
| XX | |
| DR | WPI; 2001-611632/70. |
| DR | P-BSD; AAG65767. |
| Pt | New human protein kinase polypeptides, 3714, 16742, 23546 and 13887, |
| PT | useful in diagnosis of cancer or cellular proliferation or |
| PT | differentiation disorders and to screen for polypeptide modulators useful |
| XX | to treat such conditions. |
| XX | |
| PS | Claim 1; Fig 10A-C; 169pp; English. |
| XX | |
| CC | The invention provides novel human protein kinase polypeptides, 3714, |
| CC | 16742, 23546 and 13887 and nucleic acid molecules encoding them. The |
| CC | protein kinase polypeptides can be expressed by standard recombinant |
| CC | methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides |
| CC | are useful for diagnostic and screening methods to identify subjects at |
| CC | risk of having cancer or cellular proliferation and/or differentiation |
| CC | disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and |
| CC | modulators are useful for the treatment of cancer, particularly colon |
| CC | cancer or cellular proliferation and/or differentiation disorders. Other |
| CC | disorders associated with 3714, 16742, 23546 or 13887 expression or |
| CC | activity that can be treated include bone related disorders, inflammatory |
| CC | disorders, autoimmune diseases, cardiovascular disorders and liver |
| CC | diseases. The present sequence represents a human protein kinase |
| XX | polypeptide 13887 coding sequence |
| SQ | |
| | Sequence 1260 BP, 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other; |
| | |
| | Query Match 99.6%; Score 1255.2; DB 4; Length 1260; |
| | Best Local Similarity 99.8%; Pred. No. 0; |
| | Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0 |
| OY | 1 ATGGCGGATCCGCTGCCTCGCGGGCGGAGCCGCCGTTCGTGAGGCTTCCGG 60 |
| Db | 1 ATGGCGGATCCGCTGCCTCGCGGGCGGAGCCGCCGTTCGTGAGGCTTCCGG 60 |
| OY | CGGCTGACGAGTACAAGCCGTTTGAGTAAGCGGCTCTCCGCTCGGTGATGGGTT 120 |
| Db | CGGCTGACGAGTACAAGCCGTTTGAGTAAGCGGCTCTCTCCGCTCGGTGATGGGTT 120 |
| OY | CGCTGCTCGCAACCTTGCTCGCCCCCGGCGCCTCAGGACATTCTTGCCGACAGA 180 |
| Db | CGCTGCTCGCAACCTTGCTCGCCCCCGGCGCCTCAGGACATTCTTGCCGACAGA 180 |
| OY | ACCAACCGGGCTGCGGCTCTGCGCGCGAGTAGTGTTCGCCAAGAAGGCGGCGCTG 240 |
| Db | ACCAACCGGGCTGCGGCTCTGCGCGCGAGTAGTGTTCGCCAAGAAGGCGGCGCTG 240 |
| OY | GAAACGTTGACAGGTCACAGAAACATCGTACTTTGTATGAGAGTGTATCAACTTT 300 |
| Db | GAAACGTTGACAGGTCACAGAAACATCGTACTTTGTATGAGAGTGTATCAACTTT 300 |
| OY | TCTCCAAATGACCATCAACGCTGTCTGTGCTTGAACCTCAGATGATGATTTCCGGA 360 |
| Db | TCTCCAAATGACCATCAACGCTGTCTGTGCTTGAACCTCAGATGATGATTTCCGGA 360 |
| OY | TTGCTCTTATAATTCAGTACCAAGGATTTTCATGATGATGATACAGCATTTGCCCGA 420 |
| Db | TTGCTCTTATAATTCAGTACCAAGGATTTTCATGATGATGATACAGCATTTGCCCGA 420 |
| OY | GATGTTTGAGGCGCTTGCTTTTCTTATCATATAGGGCTATGTCTCATGCGGACTCAA 480 |
| Db | GATGTTTGAGGCGCTTGCTTTTCTTATCATATAGGGCTATGTCTCATGCGGACTCAA 480 |

| | Accession | Gene | Chromosome | Start (kb) | End (kb) | Strand | Length (bp) | GC Content (%) | GC Skew | GC Bias | GC Bias2 | GC Bias3 | GC Bias4 | GC Bias5 | GC Bias6 | GC Bias7 | GC Bias8 | GC Bias9 | GC Bias10 | GC Bias11 | GC Bias12 | GC Bias13 | GC Bias14 | GC Bias15 | GC Bias16 | GC Bias17 | GC Bias18 | GC Bias19 | GC Bias20 | GC Bias21 | GC Bias22 | GC Bias23 | GC Bias24 | GC Bias25 | GC Bias26 | GC Bias27 | GC Bias28 | GC Bias29 | GC Bias30 | GC Bias31 | GC Bias32 | GC Bias33 | GC Bias34 | GC Bias35 | GC Bias36 | GC Bias37 | GC Bias38 | GC Bias39 | GC Bias40 | GC Bias41 | GC Bias42 | GC Bias43 | GC Bias44 | GC Bias45 | GC Bias46 | GC Bias47 | GC Bias48 | GC Bias49 | GC Bias50 | GC Bias51 | GC Bias52 | GC Bias53 | GC Bias54 | GC Bias55 | GC Bias56 | GC Bias57 | GC Bias58 | GC Bias59 | GC Bias60 | GC Bias61 | GC Bias62 | GC Bias63 | GC Bias64 | GC Bias65 | GC Bias66 | GC Bias67 | GC Bias68 | GC Bias69 | GC Bias70 | GC Bias71 | GC Bias72 | GC Bias73 | GC Bias74 | GC Bias75 | GC Bias76 | GC Bias77 | GC Bias78 | GC Bias79 | GC Bias80 | GC Bias81 | GC Bias82 | GC Bias83 | GC Bias84 | GC Bias85 | GC Bias86 | GC Bias87 | GC Bias88 | GC Bias89 | GC Bias90 | GC Bias91 | GC Bias92 | GC Bias93 | GC Bias94 | GC Bias95 | GC Bias96 | GC Bias97 | GC Bias98 | GC Bias99 | GC Bias100 | GC Bias101 | GC Bias102 | GC Bias103 | GC Bias104 | GC Bias105 | GC Bias106 | GC Bias107 | GC Bias108 | GC Bias109 | GC Bias110 | GC Bias111 | GC Bias112 | GC Bias113 | GC Bias114 | GC Bias115 | GC Bias116 | GC Bias117 | GC Bias118 | GC Bias119 | GC Bias120 | GC Bias121 | GC Bias122 | GC Bias123 | GC Bias124 | GC Bias125 | GC Bias126 | GC Bias127 | GC Bias128 | GC Bias129 | GC Bias130 | GC Bias131 | GC Bias132 | GC Bias133 | GC Bias134 | GC Bias135 | GC Bias136 | GC Bias137 | GC Bias138 | GC Bias139 | GC Bias140 | GC Bias141 | GC Bias142 | GC Bias143 | GC Bias144 | GC Bias145 | GC Bias146 | GC Bias147 | GC Bias148 | GC Bias149 | GC Bias150 | GC Bias151 | GC Bias152 | GC Bias153 | GC Bias154 | GC Bias155 | GC Bias156 | GC Bias157 | GC Bias158 | GC Bias159 | GC Bias160 | GC Bias161 | GC Bias162 | GC Bias163 | GC Bias164 | GC Bias165 | GC Bias166 | GC Bias167 | GC Bias168 | GC Bias169 | GC Bias170 | GC Bias171 | GC Bias172 | GC Bias173 | GC Bias174 | GC Bias175 | GC Bias176 | GC Bias177 | GC Bias178 | GC Bias179 | GC Bias180 | GC Bias181 | GC Bias182 | GC Bias183 | GC Bias184 | GC Bias185 | GC Bias186 | GC Bias187 | GC Bias188 | GC Bias189 | GC Bias190 | GC Bias191 | GC Bias192 | GC Bias193 | GC Bias194 | GC Bias195 | GC Bias196 | GC Bias197 | GC Bias198 | GC Bias199 | GC Bias200 | GC Bias201 | GC Bias202 | GC Bias203 | GC Bias204 | GC Bias205 | GC Bias206 | GC Bias207 | GC Bias208 | GC Bias209 | GC Bias210 | GC Bias211 | GC Bias212 | GC Bias213 | GC Bias214 | GC Bias215 | GC Bias216 | GC Bias217 | GC Bias218 | GC Bias219 | GC Bias220 | GC Bias221 | GC Bias222 | GC Bias223 | GC Bias224 | GC Bias225 | GC Bias226 | GC Bias227 | GC Bias228 | GC Bias229 | GC Bias230 | GC Bias231 | GC Bias232 | GC Bias233 | GC Bias234 | GC Bias235 | GC Bias236 | GC Bias237 | GC Bias238 | GC Bias239 | GC Bias240 | GC Bias241 | GC Bias242 | GC Bias243 | GC Bias244 | GC Bias245 | GC Bias246 | GC Bias247 | GC Bias248 | GC Bias249 | GC Bias250 | GC Bias251 | GC Bias252 | GC Bias253 | GC Bias254 | GC Bias255 | GC Bias256 | GC Bias257 | GC Bias258 | GC Bias259 | GC Bias260 | GC Bias261 | GC Bias262 | GC Bias263 | GC Bias264 | GC Bias265 | GC Bias266 | GC Bias267 | GC Bias268 | GC Bias269 | GC Bias270 | GC Bias271 | GC Bias272 | GC Bias273 | GC Bias274 | GC Bias275 | GC Bias276 | GC Bias277 | GC Bias278 | GC Bias279 | GC Bias280 | GC Bias281 | GC Bias282 | GC Bias283 | GC Bias284 | GC Bias285 | GC Bias286 | GC Bias287 | GC Bias288 | GC Bias289 | GC Bias290 | GC Bias291 | GC Bias292 | GC Bias293 | GC Bias294 | GC Bias295 | GC Bias296 | GC Bias297 | GC Bias298 | GC Bias299 | GC Bias300 | GC Bias301 | GC Bias302 | GC Bias303 | GC Bias304 | GC Bias305 | GC Bias306 | GC Bias307 | GC Bias308 | GC Bias309 | GC Bias310 | GC Bias311 | GC Bias312 | GC Bias313 | GC Bias314 | GC Bias315 | GC Bias316 | GC Bias317 | GC Bias318 | GC Bias319 | GC Bias320 | GC Bias321 | GC Bias322 | GC Bias323 | GC Bias324 | GC Bias325 | GC Bias326 | GC Bias327 | GC Bias328 | GC Bias329 | GC Bias330 | GC Bias331 | GC Bias332 | GC Bias333 | GC Bias334 | GC Bias335 | GC Bias336 | GC Bias337 | GC Bias338 | GC Bias339 | GC Bias340 | GC Bias341 | GC Bias342 | GC Bias343 | GC Bias344 | GC Bias345 | GC Bias346 | GC Bias347 | GC Bias348 | GC Bias349 | GC Bias350 | GC Bias351 | GC Bias352 | GC Bias353 | GC Bias354 | GC Bias355 | GC Bias356 | GC Bias357 | GC Bias358 | GC Bias359 | GC Bias360 | GC Bias361 | GC Bias362 | GC Bias363 | GC Bias364 | GC Bias365 | GC Bias366 | GC Bias367 | GC Bias368 | GC Bias369 | GC Bias370 | GC Bias371 | GC Bias372 | GC Bias373 | GC Bias374 | GC Bias375 | GC Bias376 | GC Bias377 | GC Bias378 | GC Bias379 | GC Bias380 | GC Bias381 | GC Bias382 | GC Bias383 | GC Bias384 | GC Bias385 | GC Bias386 | GC Bias387 | GC Bias388 | GC Bias389 | GC Bias390 | GC Bias391 | GC Bias392 | GC Bias393 | GC Bias394 | GC Bias395 | GC Bias396 | GC Bias397 | GC Bias398 | GC Bias399 | GC Bias400 | GC Bias401 | GC Bias402 | GC Bias403 | GC Bias404 | GC Bias405 | GC Bias406 | GC Bias407 | GC Bias408 | GC Bias409 | GC Bias410 | GC Bias411 | GC Bias412 | |
|--|-----------|------|------------|------------|----------|--------|-------------|----------------|---------|---------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--|
|--|-----------|------|------------|------------|----------|--------|-------------|----------------|---------|---------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--|

```

FT      CDS      /*tag= a
FT      267..1529 /*tag= b
FT      /product= "protein kinase 13887"
FT      /note= "coding sequence specifically claimed"
FT      1530..2622
FT      3'UTR     /*tag= c
FT
XX      MO200173050-A2.
XX
XX      04-OCT-2001.
XX
XX      23-MAR-2001; 2001MO-US009483.
XX
XX      24-MAR-2000; 2000US-0191846P.
XX
XX      (MILL-) MILLENNIUM PHARM INC.
XX
XX      Meyers R;
XX
XX      MPI; 2001-611632/70.
XX      P-PSDB; AAG65767.
XX
XX      New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
XX      useful in diagnosis of cancer or cellular proliferation or
XX      differentiation disorders and to screen for polypeptide modulators useful
XX      to treat such conditions.
XX
XX      Claim 1; Fig 10A-C; 169pp; English.
XX
XX      The invention provides novel human protein kinase polypeptides, 3714,
XX      16742, 23546 and 13887 and nucleic acid molecules encoding them. The
XX      protein kinase polypeptides can be expressed by standard recombinant
XX      methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
XX      are useful for diagnostic and screening methods to identify subjects (at
XX      risk of) having cancer or cellular proliferation and/or differentiation
XX      disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
XX      modulators are useful for the treatment of cancer, particularly colon
XX      cancer or cellular proliferation and/or differentiation disorders. Other
XX      disorders associated with 3714, 16742, 23546 or 13887 expression or
XX      activity that can be treated include bone related disorders, inflammatory
XX      disorders, autoimmune diseases, cardiovascular disorders and liver
XX      diseases. The present sequence represents a human protein kinase
XX      polypeptide 13887 encoding cDNA
XX
XX      Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;
SQ
Query Match      99.6%; Score 1255.2; DB 4; Length 2622;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Db      567 TCTCCAAATGTCATCAACGCTGTCTGTTGAACTTCGATGTCACTGTTTCGGA 626
Qy      361 TTGCTCTTATATTCAGTCACAGGAGTGTTCATATGGAGATATCAGCATTTGCCCGGA 420
Db      627 TTGCTCTTATATTCAGTCACAGGAGTGTTCATATGGAGATATCAGCATTTGTGCCGA 686
Qy      421 GATGTTTGGAGCCCTTGCTTTCTTCATCATGAGGCTATGTCCATGCGGACTCAAA 480
Db      687 GATGTTTGGAGCCCTTGCTTTCTTCATCATGAGGCTATGTCCATGCGGACTCAAA 746
Qy      481 CCACGTAAATATGAGAGTCAGAGATGAATGTTTAACTCATGACTTTGGAAGT 540
Db      747 CCACGTAAATATGAGAGTCAGAGATGAATGTTTAACTCATGACTTTGGAAGT 806
Qy      541 AGCTTCAAGAGAGCAATCAGATGTAAAGTATATTCAGACAGACGGGATCGGCTCCA 600
Db      807 AGCTTCAAGAGAGCAATCAGATGTAAAGTATATTCAGACAGACGGGATCGGCTCCA 866
Qy      601 GAAGCAAGATTGCAAAATTTGCTTGCCAGGCTGGCTGCAAGATGATCAGATGTAC 660
Db      867 GAAGCAAGATTGCAAAATTTGCTTGCCAGGCTGGCTGCAAGATGATCAGATGTAC 926
Qy      661 TCAGCTGTTGATCTGTGAGAGCTAGAGAAATCATTTTACTGGAATGTTCTCAGAAATGAA 720
Db      927 TCAGCTGTTGATCTGTGAGAGCTAGAGAAATCATTTTACTGGAATGTTCTCAGAAATGAA 986
Qy      721 CTGAAACATACAGTCAGATCTCAGAGATGGAAGCAACAGTTCTGATTAATGATCAC 780
Db      987 CTGAAACATACAGTCAGATCTCAGAGATGGAAGCAACAGTTCTGATTAATGATCAC 1046
Qy      781 ATATTTGCCAGTAAAGCAATGAGTGAATGCCCAATTCAGCTTATCACTTAAGACCTT 840
Db      1047 ATATTTGCCAGTAAAGCAATGAGTGAATGCCCAATTCAGCTTATCACTTAAGACCTT 1106
Qy      841 ATCAAAAGCATGCTTCATGATGATCAAGCAGAGAAATTCCTGTGAAATGACATTTGTC 900
Db      1107 ATCAAAAGCATGCTTCATGATGATCAAGCAGAGAAATTCCTGTGAAATGACATTTGTC 1166
Qy      901 AGCCCATTTTAAAGATTCCTTTTGCCTCATATTTGAAGATCTGATGCTTCCACT 960
Db      1167 AGCCCATTTTAAAGATTCCTTTTGCCTCATATTTGAAGATCTGATGCTTCCACT 1226
Qy      961 CCAGTCTTAAGCTGCTGAATGCTGATGATGATTAATCTTTGGAAATGGAAGAAAT 1020
Db      1227 CCAGTCTTAAGCTGCTGAATGCTGATGATGATTAATCTTTGGAAATGGAAGAAAT 1286
Qy      1021 GAAGATGTTGAAGATGTAAGAGAGAGTCAAAAATATGACCAAGTGTATCTCTA 1080
Db      1287 GAAGATGTTGAAGATGTAAGAGAGAGTCAAAAATATGACCAAGTGTATCTCTA 1346
Qy      1081 CTTGTTCCAAAGAAATCTTGCAAGAGCAAGTCTTTGTTGAGTATGCAAAATCTGGT 1140
Db      1347 CTTGTTCCAAAGAAATCTTGCAAGAGCAAGTCTTTGTTGAGTATGCAAAATCTGGT 1406
Qy      1141 GATTCGAAGGCTGCCAGAAATTAATGACTGAAGAATGTTTGAATGGAAATTTTGTG 1200
Db      1407 GATTCGAAGGCTGCCAGAAATTAATGACTGAAGAATGTTTGAATGGAAATTTTGTG 1466
Qy      1201 GCTACATTTACCCGCTGAGTCTTACAGAGAGGATATCTGTATCAAACTTTGCTTTAA 1260
Db      1467 GCTACATTTACCCGCTGAGTCTTACAGAGAGGATATCTGTATCAAACTTTGCTTTAA 1526

```

```

RESULT 6
AAH72714
ID      AAH72714 Standard; cDNA; 4065 BP.
XX
XX      AAH72714;
XX
XX      19-SEP-2001 (first entry)
XX
XX      Human cervical cancer marker nucleic acid 3988.
XX

```


| | | | |
|--|------|---|------|
| Db | 688 | GAGTTTGGAGGCCCTTGCTTTCTTCATCAGAGGCGTAATGCATGGGACCTGAAA | 747 |
| Qy | 481 | CCACGTAAACAATTTGGAGTGCAGAGATGATGTTTAACTCATTTGACTTTGGACTT | 540 |
| Db | 748 | CCACGTAAACAATTTGGAGTGCAGAGATGATGTTTAACTCATTTGACTTTGGACTT | 807 |
| Qy | 541 | AGCTTCAAAGAAGGCAATCAGAGATGTAAGTAACTTAATTCAGACAGCGGGTATCGGGCTCCA | 600 |
| Db | 808 | AGCTTCAAAGAAGGCAATCAGAGATGTAAGTAACTTAATTCAGACAGCGGGTATCGGGCTCCA | 867 |
| Qy | 601 | GAAGCAGAAATTTGCAAAATTTGCTTTGGCCGAGGCTGGCTGCAAGTGTATACAGATGTACC | 660 |
| Db | 868 | GAAGCAGAAATTTGCAAAATTTGCTTTGGCCGAGGCTGGCTGCAAGTGTATACAGATGTACC | 927 |
| Qy | 661 | TCAGCTGTGATCTGTGGAGCTTAGAATCATTTTACTGAAATGTTCTCAGAAATGAAA | 720 |
| Db | 928 | TCAGCTGTGATCTGTGGAGCTTAGAATCATTTTACTGAAATGTTCTCAGAAATGAAA | 987 |
| Qy | 721 | CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAACAGTTCTGTATTAATGATCAC | 780 |
| Db | 988 | CTGAAACATACAGTCAGATCTCAGGAATGGAAGGCAACAGTTCTGTATTAATGATCAC | 1047 |
| Qy | 781 | ATATTTCCAGTAAAGAGTGTGAATGCGGCAATTCAGGCTTTCACCTAAGAGACTT | 840 |
| Db | 1048 | ATATTTCCAGTAAAGAGTGTGAATGCGGCAATTCAGGCTTTCACCTAAGAGACTT | 1107 |
| Qy | 841 | ATCAAAGAATGCTTCAATGATATCCAAAGCAAGAAATTCCTGTGAAATGGCATTTGTC | 900 |
| Db | 1108 | ATCAAAGAATGCTTCAATGATATCCAAAGCAAGAAATTCCTGTGAAATGGCATTTGTC | 1167 |
| Qy | 901 | AGCCATTCTTTAGCATTCCTTTTGCCCTTCATATGGAAGTGTGATGCTTCCACT | 960 |
| Db | 1168 | AGCCATTCTTTAGCATTCCTTTTGCCCTTCATATGGAAGTGTGATGCTTCCACT | 1227 |
| Qy | 961 | CCAGTGTAAAGCTGCGGAAATGTCGAGATGATGATATCTGTGGATGAAGAGGAATAT | 1020 |
| Db | 1228 | CCAGTGTAAAGCTGCGGAAATGTCGAGATGATGATATCTGTGGATGAAGAGGAATAT | 1287 |
| Qy | 1021 | GAAGATTTGTAGAGATGTAAAGAAGAGTGTCAAAAATATGGAACCACTGTATCTCTA | 1080 |
| Db | 1288 | GAAGATTTGTAGAGATGTAAAGAAGAGTGTCAAAAATATGGAACCACTGTATCTCTA | 1347 |
| Qy | 1081 | CTTGTTCCAAAGAAAATCCTGGCAGAGACAAAGTCTTTGTGAGTATGCAAAATGCTGGT | 1140 |
| Db | 1348 | CTTGTTCCAAAGAAAATCCTGGCAGAGACAAAGTCTTTGTGAGTATGCAAAATGCTGGT | 1407 |
| Qy | 1141 | GATTTCCAAAGCTGCGCAAAATTAATCTGACTGGAAGATGTTTGAATGGAGTTTGTGTG | 1200 |
| Db | 1408 | GATTTCCAAAGCTGCGCAAAATTAATCTGACTGGAAGATGTTTGAATGGAGTTTGTGTG | 1467 |
| Qy | 1201 | GCTACATTTTACCGGCTGAGTGCCTACAAAGAGGGAATATCTGTATCAAACTTGTCTTTAA | 1260 |
| Db | 1468 | GCTACATTTTACCGGCTGAGTGCCTACAAAGAGGGAATATCTGTATCAAACTTGTCTTTAA | 1527 |
| RESULT 7 | | | |
| ADL45238 | | | |
| ID ADL45238 standard; DNA: 4065 BP. | | | |
| ADL45238; | | | |
| 20-MAY-2004 (first entry) | | | |
| Human ovarian cancer DNA marker #19128. | | | |
| Human; ovarian cancer; ds; tumour; cytostatic; DNA marker. | | | |
| Homo sapiens. | | | |
| W0200170979-A2. | | | |
| 27-SEP-2001. | | | |

AA06734 standard; cDNA; 1260 BP.
AA06734;
12-SEP-2001 (first entry)
Polynucleotide sequence encoding human protein kinase #34.
Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
metabolic disorder; immune related disease; neurological disorder;
neurodegenerative disorder; inflammatory disorder; infectious disease;
reproductive disorder; gene therapy; ss.
Homo sapiens.
WO200138503-A2.
31-MAY-2001.
22-NOV-2000; 2000MO-US032085.
24-NOV-1999; 99US-0167482P.
(SUGR-) SUGEN INC.
Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
Flanagan P, Clary D;
WPI; 2001-343950/36.
P-PSDB; AAU03534.
Nucleic acids encoding human kinase polypeptides, useful for preventing
diagnosing and/or treating e.g. cancer, immune, cardiovascular and
neural-associated diseases, and microbial infections.
Example 1; Fig 1; 433bp; English.
AA06701-AA06757 encode for novel human protein kinases #1-57. The novel
protein kinases have been identified as members of the tyrosine or
serine/threonine kinase (PTK and STK) families. The polynucleotides
encoding protein kinases and the polypeptides may be used in the
prevention, diagnosis and treatment of diseases associated with
inappropriate kinase expression. For example, they may be used to treat
cancers (especially cancers of hematopoietic origin), cardiovascular
disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
immune related diseases (e.g. rheumatoid arthritis), neurological
disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
disease (e.g. HIV) and reproductive disorders (e.g. infertility).
Additionally, polynucleotides encoding protein kinases may be used for
gene therapy and as DNA probes in diagnostic assays. The protein kinase
polypeptides may be used as antigens in the production of antibodies
against the protein kinases and in assays to identify modulators of
protein kinase expression and activity
Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;
Query Match 99.5%; Score 1253.6; DB 4; Length 1260;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1256; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGGCGGGAATCCGGCTGGCGCTGGGGGCGGAGCCCGCGCTTTTCGAGAGCGCTTCGGG 60
DB 1 ATGGCGGGAATCCGGCTGGCGCTGGGGGCGGAGCCCGCGCTTTTCGAGAGCGCTTCGGG 60
QY CGGCTGTGGCAGGTACAGAGCCGCTGTGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
DB 61 CGGCTGTGGCAGGTACAGAGCCGCTGTGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
QY 121 CGCTGTGGCGCAACCTGTGCTGCCCCCGGCGGCTCTCAAGGAGTTCTTGGCCGCAAGA 180
DB 121 CGCTGTGGCGCAACCTGTGCTGCCCCCGGCGGCTCTCAAGGAGTTCTTGGCCGCAAGA 180

181 ACCACCGGGGCTGCGGCTCTGCCCCGAGATAGTGTTCGCGAAGAGAGGGCGCGCTG 240
DB 181 ACCACCGGGGCTGCGGCTCTGCCCCGAGATAGTGTTCGCGAAGAGAGGGCGCGCTG 240
QY 241 GAACAGTTGACAGGTTCACAGAAACATCGTACTTTGTATGAGTGTTCACATCCATT 300
DB 241 GAACAGTTGACAGGTTCACAGAAACATCGTACTTTGTATGAGTGTTCACATCCATT 300
QY 301 TCTCCAAATGTCGATCAACGCTGTCTGTTCGTTGAATCTTCGAGATGTCAGTGTTCGGA 360
DB 301 TCTCCAAATGTCGATCAACGCTGTCTGTTCGTTGAATCTTCGAGATGTCAGTGTTCGGA 360
QY 361 TTGCTCTTATATTCAGTCAACAGAGTGTTCATGATGATATCAGATTCGCCCGCA 420
DB 361 TTGCTCTTATATTCAGTCAACAGAGTGTTCATGATGATATCAGATTCGCCCGCA 420
QY 421 GATGTTTTGAGAGCCCTTGTCTTTCATCATGAGGCTATGTCATGCGGACCTCAAA 480
DB 421 GATGTTTTGAGAGCCCTTGTCTTTCATCATGAGGCTATGTCATGCGGACCTCAAA 480
QY 481 CCAAGTAACTATTTGTGAGTGCAGAGATGAATGATTTTAACTGATTCATTTGAGCTT 540
DB 481 CCAAGTAACTATTTGTGAGTGCAGAGATGAATGATTTTAACTGATTCATTTGAGCTT 540
QY 541 AGCTTCAAAAGAGCAATCAGGATGTAAAGTATATTCAGACAGAGGATTCGGCTCCA 600
DB 541 AGCTTCAAAAGAGCAATCAGGATGTAAAGTATATTCAGACAGAGGATTCGGCTCCA 600
QY 601 GAAGCAGAAATTGCATAATTCCTTGGCCGAGCTGCGCTGAGATGATCAGATGTAC 660
DB 601 GAAGCAGAAATTGCATAATTCCTTGGCCGAGCTGCGCTGAGATGATCAGATGTAC 660
QY 661 TCAGCTGTGATCTGTGAGAGCTTGAATATTTTAACTGGAATGTTCTCAGAAATGAA 720
DB 661 TCAGCTGTGATCTGTGAGAGCTTGAATATTTTAACTGGAATGTTCTCAGAAATGAA 720
QY 721 CTGAAACATACAGTCAAGATCTCAGAAATGAAAGCAACAGTCTGTATTTATGATCAC 780
DB 721 CTGAAACATACAGTCAAGATCTCAGAAATGAAAGCAACAGTCTGTATTTATGATCAC 780
QY 781 ATATTTGCGAGTAAAGCATGTGTGAATGCGCAATTCAGCTTATCATAAGACCTT 840
DB 781 ATATTTGCGAGTAAAGCATGTGTGAATGCGCAATTCAGCTTATCATAAGACCTT 840
QY 841 ATCAAAAGATGCTTATGATATCCAGCAGAAATATTCCTGCTGAATGCAATGTC 900
DB 841 ATCAAAAGATGCTTATGATATCCAGCAGAAATATTCCTGCTGAATGCAATGTC 900
QY 901 AGCCCATTTCTTGAAGCTTCTTGGCCCTCATATTTGAAGATCTGTCACTTCCACT 960
DB 901 AGCCCATTTCTTGAAGCTTCTTGGCCCTCATATTTGAAGATCTGTCACTTCCACT 960
QY 961 CCAAGTCTAAGATCTGTGAATGTGTGATGATATCTTGGAAATGAGAGAAATAT 1020
DB 961 CCAAGTCTAAGATCTGTGAATGTGTGATGATATCTTGGAAATGAGAGAAATAT 1020
QY 1021 GAAAGATTTGTGAAGATGTAAAGAGATGTCAAAAAATATGACCACTGTATCTCA 1080
DB 1021 GAAAGATTTGTGAAGATGTAAAGAGATGTCAAAAAATATGACCACTGTATCTCA 1080
QY 1081 CTTGTTCCAAAGAAATCTGTGCAAGAGCAAGCTTGTGTGAATGCAATGCTGTGT 1140
DB 1081 CTTGTTCCAAAGAAATCTGTGCAAGAGCAAGCTTGTGTGAATGCAATGCTGTGT 1140
QY 1141 GATTCCAAAGCTGTGCAAGAAATTAAGCTGGAAGAGATGTTGATGGAAGTGTGTG 1200
DB 1141 GATTCCAAAGCTGTGCAAGAAATTAAGCTGGAAGAGATGTTGATGGAAGTGTGTG 1200
QY 1201 GCTACATTTACCGCTGATGCTCAAGAGGAGATATCTGATCAACCTTGTCTTAA 1260
DB 1201 GCTACATTTACCGCTGATGCTCAAGAGGAGATATCTGATCAACCTTGTCTTAA 1260

RESULT 9
 AAD18818
 ID AAD18818 standard; cDNA; 1824 BP.
 XX
 AC AAD18818;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human Kinase (PKIN) -3 CDNA.
 XX
 KW Human kinase; PKIN, gene therapy; adenocarcinoma; immune disorder; gout; cancer; allergy; sarcoma; leukemia; acquired immune deficiency syndrome; AIDS; Addison's disease; microbial infection; inflammation; osteoporosis; atherosclerosis; cardiovascular disease; myocardial infarction; anaemia; myasthenia gravis; cirrhosis; cataract; growth and development disorder; seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder; lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease; asthma; obesity; restorative therapy; cytostatic; immunomodulatory; antimicrobial; cardiovascular; anti-inflammatory; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 169..1428
 FT /tag= a
 FT /product= "Human PKIN-3 protein"
 XX
 PN MO200181555-A2.
 XX
 XX 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001MO-US012992.
 XX
 PR 20-APR-2000; 2000US-0199021P.
 PR 28-APR-2000; 2000US-0200226P.
 PR 05-MAY-2000; 2000US-0202339P.
 PR 11-MAY-2000; 2000US-0203505P.
 PR 18-MAY-2000; 2000US-0205564P.
 PR 26-MAY-2000; 2000US-0207739P.
 PR 01-JUN-2000; 2000US-0208795P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
 PI Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
 PI Patterson C, Burdill JD, Marcus GA, Zingler KA, Reardon SP, Lu Y;
 PI Pollocky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
 PI Walsh RT, Rankumar J, Borowsky ML, Au-Young J, Hillman JL;
 PI Gururajan R;
 XX
 DR MPI; 2001-611740/70.
 DR P-PSDB; AAE11769.
 XX
 PT Human kinases and nucleic acids, useful for preventing diagnosing and
 PT treating cancers, inflammation and immune disorders.
 XX
 PS Claim 5; Page 152; 166pp; English.
 XX
 CC The present invention relates to human kinases (PKIN) and the nucleic
 CC acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
 CC used in the prevention, diagnosis and treatment of diseases cancers,
 CC adenocarcinoma, leukemia, sarcoma, immune disorder, Addison's disease,
 CC acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
 CC gout, microbial infections, cardiovascular disease and/or inflammation,
 CC myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
 CC infarction, cataract, growth and development disorder, seizure disorder,
 CC pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
 CC disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
 CC PKIN may be used to treat disorders associated with decreased PKIN
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of PKIN by expressing inactive proteins or to
 CC supplement the patients own production of PKIN. PKIN nucleic acids may be
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into

CC a host cell and culturing the cell to express the protein. PKIN nucleic
 CC acid and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays to detect and quantitate the presence of similar
 CC nucleic acid sequences in samples and therefore which patients may be in
 CC need of restorative therapy. The present sequence is human PKIN-3 CDNA
 XX
 SQ Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;
 Query Match 99.5%; Score 1253.6; DB 4; Length 1824;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1256; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ATGGCGGGATCCGCTGGCGCTGGGGCGGAGCCGCCGCTTTCTGAGAGCCTTCGGG 60
 DB 169 ATGGCGGGATCCGCTGGCGCTGGGGCGGAGCCGCCGCTTTCTGAGAGCCTTCGGG 228
 QY 61 CGGCTGTGGCAGGTACAGAGCCGCTGSGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 120
 DB 229 CGGCTGTGGCAGGTACAGAGCCGCTGSGGTAGCGGCTCTCCGCTCGGTGTATCGGGTT 288
 QY 121 CGTGTGGCGGACCTTGCTGCGCCGCCCGCGGCTTCAGGACGTTCTTGGCCGACAGA 180
 DB 289 CGTGTGGCGGACCTTGCTGCGCCGCCCGCGGCTTCAGGACGTTCTTGGCCGACAGA 348
 QY 181 ACCACCGGGGCTGGCGGCTTGTGCGCGGAGTATGTTTCCGAAAGAGGGCGGCGCTG 240
 DB 349 ACCACCGGGGCTGGCGGCTTGTGCGCGGAGTATGTTTCCGAAAGAGGGCGGCGCTG 408
 QY 241 GAACAGTTGCAGGGTACAGAAACATCGTACTTTGTATGAGGTATTAATCCACTTT 300
 DB 409 GAACAGTTGCAGGGTACAGAAACATCGTACTTTGTATGAGGTATTAATCCACTTT 468
 QY 301 TCTCCAAATGTGCATACAGCTGTGTGCTTGAATCTTGATATGATGTTTCGGA 360
 DB 469 TCTCCAAATGTGCATACAGCTGTGTGCTTGAATCTTGATATGATGTTTCGGA 528
 QY 361 TTGCTCTTATTTCCAGTACAGGAGTTGTTCAATGATGATATCAGCATTTGGCCCGGA 420
 DB 529 TTGCTCTTATTTCCAGTACAGGAGTTGTTCAATGATGATATCAGCATTTGGCCCGGA 588
 QY 421 GATGTTTGGAGGCGCTTCTTCTCATCATGAGGGGTATGTCATGCGGACCTCAAA 480
 DB 589 GATGTTTGGAGGCGCTTCTTCTCATCATGAGGGGTATGTCATGCGGACCTCAAA 648
 QY 481 CCACGTAATATTTGAGAGTGCAGAGATGATGATTTTAACTGATGACTTTGACTT 540
 DB 649 CCACGTAATATTTGAGAGTGCAGAGATGATGATTTTAACTGATGACTTTGACTT 708
 QY 541 AGCTTCAAGAAAGCAATCAGATGTAAAGTATTTACAGACAGAGGGTATCGGCTCCA 600
 DB 709 AGCTTCAAGAAAGCAATCAGATGTAAAGTATTTACAGACAGAGGGTATCGGCTCCA 768
 QY 601 GAAGCAGATTTGCATAATTTGCTTGGCCAGGCTGGCTGCAGAGTATCAGATGTACC 660
 DB 769 GAAGCAGATTTGCATAATTTGCTTGGCCAGGCTGGCTGCAGAGTATCAGATGTACC 828
 QY 661 TCAGCTGTGATCTGTGAGGCTTGAAGATCAATTTTACTGAAATGTTCTCAGAAATGAA 720
 DB 829 TCAGCTGTGATCTGTGAGGCTTGAAGATCAATTTTACTGAAATGTTCTCAGAAATGAA 888
 QY 721 CTGAAACATACGTAGATCTCAGAAATGAAAGCAACAGTTCTGCTATTTATATCAC 780
 DB 889 CTGAAACATACGTAGATCTCAGAAATGAAAGCAACAGTTCTGCTATTTATATCAC 948
 QY 781 AATTTTGCAGTAAGAGGTGTGAATCGCGAATTTCCAGCTTATCAGCTTAAGAGACCTT 840
 DB 949 AATTTTGCAGTAAGAGGTGTGAATCGCGAATTTCCAGCTTATCAGCTTAAGAGACCTT 1008
 QY 841 ATCAAAAGCATCTTCATGATGATCCAAAGCAGAAAGATTTCTGTAAGATGCAATTTGTC 900
 DB 1009 ATCAAAAGCATCTTCATGATGATCCAAAGCAGAAAGATTTCTGTAAGATGCAATTTGTC 1068
 QY 901 AGCCCATCTTATGACATTCCTTTTGGCCCTCATTTATGAAGATGTGTCATGCTTCCACT 960


```
QY 721 CTGAAACATACAGTCAGATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATTATGATCAC 780
DB 882 CTGAAACATACAGTCAGATCTCAGAAATGGAAGCAAAAGTTCTGCTATTATTATGATCAC 941
QY 781 ATATTTTCCAGTAAAGAGAGTGTGAATGCCGCAATTCAGCTTATCATCCTAAGAGACCTT 840
DB 942 ATATTTTCCAGTAAAGAGAGTGTGAATGCCGCAATTCAGCTTATCATCCTAAGAGACCTT 1001
QY 841 ATCAAAAGCATGCTTATGATGATCCAGAGCAAGAAATCCCTGCTAAATGAGATGTGC 900
DB 1002 ATCAAAAGCATGCTTATGATGATCCAGAGCAAGAAATCCCTGCTAAATGAGATGTGC 1061
QY 901 AGCCCATTTCTTTAGCATTCCTTTTGCCTCATATTGAAGATGTGGTCAATGCTTCCACT 960
DB 1062 AGCCCATTTCTTTAGCATTCCTTTTGCCTCATATTGAAGATGTGGTCAATGCTTCCACT 1121
QY 961 CCAGTCTCAAGACTGCTGAATGTGCTGATGATGATTAATCTTGGAAATGAGAAATAT 1020
DB 1122 CCAGTCTCAAGACTGCTGAATGTGCTGATGATGATTAATCTTGGAAATGAGAAATAT 1181
QY 1021 GAAGATGTGTAAAGATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTA 1080
DB 1182 GAAGATGTGTAAAGATGTAAAGAGAGTGTCAAAAATATGACCAAGTGTATCTCTG 1241
QY 1081 CTGTGTCCAAAGGAAATCCCTGGCAGAGCAAGTCTTTGTGATGCAAAATGCTGGT 1140
DB 1242 CTGTGTCCAAAGGAAATCCCTGGCAGAGCAAGTCTTTGTGATGCAAAATGCTGGT 1301
QY 1141 GATTCCAAAGCTGCGGAGAAATTAATCTGACTGGAAGATGTTTGATGGAAATGTTGTG 1200
DB 1302 GATTCCAAAGCTGCGGAGAAATTAATCTGACTGGAAGATGTTTGATGGAAATGTTGTG 1361
QY 1201 GCTACATTTCTACCCGCTGAGTGTCTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1260
DB 1362 GCTACATTTCTACCCGCTGAGTGTCTACAGAGGGGATATCTGTATCAAACTTGTCTTAA 1421
```

RESULT 12

AAH70764
ID AAH70764 standard; cDNA; 529 BP.

AC AAH70764;
XX
DT 19-SEP-2001 (first entry)

DE Human cervical cancer marker nucleic acid 2038.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

OS Homo sapiens.

PN WO200142467-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-US033312.

PR 08-DEC-1999; 99US-0169681P.

PR 21-DEC-1999; 99US-0171350P.

PR 14-MAR-2000; 2000US-0189315P.

PR 12-MAY-2000; 2000US-0203791P.

PR 09-JUN-2000; 2000US-0210600P.

PR 21-JUL-2000; 2000US-0220114P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Deeds J, Berger A, Zhao X;

XX WPI; 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer and
for assessing and detecting compounds for treating the cancer.

XX
PS Claim 1; Page 434; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX

Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Query Match 36.4%; Score 458.6; DB 4; Length 529;

Best Local Similarity 97.8%; Pred. No. 7,1e-126;

Matches 486; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 654 ATGTACTCAGCTGTGATCTGTGAGAGCTAGAGATCAATTTTACTGAAATGTTCTCAGG 713

DB 22 AGTACCTCAGCTGTGATCTGTGAGAGCTAGAGATCAATTTTACTGAAATGTTCTCAGG 81

QY 714 AATGAACTGAAACATACAGTCAGATCTCAGGAATGAAAGCAACGTTCTGTATTAT 773

DB 82 AATGAACTGAAACATACAGTCAGATCTCAGGAATGAAAGCAACGTTCTGTATTAT 141

QY 774 TGATCAGATTTTGGCAGTAAAGCAGTGTGAATGCCCAATTCAGCTTACCTTAAG 833

DB 142 TGATCAGATTTTGGCAGTAAAGCAGTGTGAATGCCCAATTCAGCTTACCTTAAG 201

QY 834 ACACCTTATCAAAAGCATGCTCATGATGATCAAGAGAAATTCCTGTAATGGC 893

DB 202 ACACCTTATCAAAAGCATGCTCATGATGATCAAGAGAAATTCCTGTAATGGC 261

QY 894 ATGTGACAGCCCATTCCTTACGATTCCTTTGCCCCCTCATATTGAAGATCTGTATGCT 953

DB 262 ATGTGACAGCCCATTCCTTACGATTCCTTTGCCCCCTCATATTGAAGATCTGTATGCT 321

QY 954 TCCCACTCAGTGTGAAGATCTGTAATGTGTGATGATGATTAATCTTGGAAATGAAGA 1013

DB 322 TCCCACTCAGTGTGAAGATCTGTAATGTGTGATGATGATTAATCTTGGAAATGAAGA 381

QY 1014 GGATATGAGA- TGTGTGAGAAGATGTAAGA- GAGTGTCAAAAATATGAGACAGTG 1071

DB 382 GGATATGAGAAGTGTGTGAGAAGATGTAAGAAGAGAGTGTCAAAAATATGAGACAGTG 441

QY 1072 GTATCTACTTGTTCAAAGGAAATCCTGGCAGAGGACAGTCTTGTGATGATGCA 1131

DB 442 GTATCTACTTGTTCAAAGGAAATCCTGGCAGAGGACAGTCTTGTGATGATGCA 501

QY 1132 AATGCTGTGATTCGA 1148

DB 502 AAGCTGGGATTCGAA 518

RESULT 13

AAA02535
ID AAA02535 standard; cDNA; 722 BP.

AC AAA02535;

DT 19-MAY-2000 (first entry)

DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.

XX Human; colon cancer; tumour; diagnosis; gene expression product; probe;

XX detection; cancerous state; metastasis; identification; breast cancer;

XX oestrogen receptor-negative breast cancer; lung cancer; ss.

OS Homo sapiens.

PN WO9508675-A2.

XX 18-NOV-1999.
 PD 13-MAY-1999; 99WO-US010602.
 XX 14-MAY-1999; 98US-0085426P.
 PR 15-MAY-1998; 98US-0085537P.
 PR 15-MAY-1998; 98US-0085696P.
 PR 21-OCT-1998; 98US-0105234P.
 PR 27-OCT-1998; 98US-0105877P.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Gleese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 PI Lamsan G, Drmanac R, Ckvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones LW, Strache-Crain B;
 XX WPI, 2000-126369/11.
 DR
 XX Polynucleotide library used to determine cancerous states of mammalian
 PT cells.
 XX Claim 1, Page 1020, 1097PP; English.

XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
 CC libraries constructed from human colon cancer cell lines. The present
 CC invention also describes a method of detecting differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test
 CC sample derived from a cell suspected of being cancerous, where detection
 CC of the differentially expressed gene product is correlated with a
 CC cancerous state of the cell from which the test sample was derived. The
 CC polynucleotide sequences can be used in a method for detecting
 CC differentially expressed genes correlated with a cancerous state of a
 CC mammalian cell. The polynucleotides can also be used as probes for
 CC detecting and mapping related genes. They can be used in diagnosis and
 CC prognosis of diseases and disorders (e.g. identification of pre-
 CC metastatic or metastatic cancerous states, stages of cancer, or
 CC responsiveness of cancer to therapy). This is particularly for breast
 CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
 CC negative breast cancer, lung cancer, and colon cancer
 CC
 XX
 SQ Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

Query Match 35.4%; Score 445.8; DB 3; Length 722;
 Best Local Similarity 93.0%; Pred. No. 5.6e-122; Indels 15; Gaps 11;
 Matches 581; Conservative 0; Mismatches 29;

QY 454 GAGGGCTATGTCCATCGGACCTCAACACGTAATATTGTGAGTCAGAGAAATGAA 513
 DB 1 GAGGGCTATGTCCATCGGACCTCAACACGTAATATTGTGAGTCAGAGAAATGAA 60
 QY 514 TGTTTTAACTCAATTGACTTTGACCTTCAAGAGGCAATCAAGATGTAAGTAT 573
 DB 61 TGTTTTAACTCAATTGACTTTGACCTTCAAGAGGCAATCAAGATGTAAGTAT 120
 QY 574 ATTCAACACGAGGGGTATCGGGCTCCAGACGAAATTTGCAAAATGCTTGGCCAGGCT 633
 DB 121 ATTCAACACGAGGGGTATCGGGCTCCAGAA-CAGAAATGCAAAATGCTTGGCCAGGCT 179
 QY 634 GGCCTCAGAGTATACAGAAATGACTTGTGATCTGTGAGCCTTGAAGATCAT 693
 DB 180 GGCCTCAGAGTATACAGAAATGACTTGTGATCTGTGAGCCTTGAAGATCAT 239
 QY 694 TTAATCGAAATGTTCTCAGAAATGAAATGAAATCAATACAGTCAATCTCAGGAATGAG 753
 DB 240 TTAATCGAAATGTTCTCAGAAATGAAATGAAATCAATACAGTCAATCTCAGGAATGAG 299
 QY 754 GCAAAAG-TTTCGCTATTA-TTGATCAATATTTGCAATAA--GCAGTGGTGAATGC 809
 DB 300 GCAAAAGTTTCTGCTATTAATTTGATCAATATTTGCAATAAAGCAANTGTTGAATGC 359

QY 810 CGCAATTCAGGCTATTCACCTTAAGACCTTATCAAAAGCATGCTTCAGATGATCC--A 867
 DB 360 CGCAATTCAGGCTATTCACCTTAAGACCTTATCAAAAGCATGCTTCAGATGATCCAA 419
 QY 868 ACAGAGAAATCCCGTGAATATGCGATT-GTGACGCCATCTCTTACG-ATTCTTTTG 925
 DB 420 GCAGAGAAATTTCTNCTGAATATGCGATTGTGTCANCCATCTTTAGCAATCTTTTG 479
 QY 926 CCCCTCATATTTGAAGATCTGGTATGCTTCCATCTCCAGT-GCTAAGACTGCTGAATGTG 984
 DB 480 CCCCTCATATTTGAAGATCTGGTATGCTTCCATCTCCAGTGGCTAAGACTGCTGAATGTG 539
 QY 985 CT-GGATGATGTTATCTTTGGGAATGAA--GAGGAATATGAAGATGTTGAAGATGT- 1040
 DB 540 CTGGGATGATGTTATCTTTGAAGATGAAAGAGATTTGAAGATGTTGTAAGATGNT 599
 QY 1041 -AAAAGAGAGTGTCAAAATATATGG 1064
 DB 600 AAAAGAGAGTGTGCAAAATATTTTG 624

RESULT 14

AAH71966
 ID AAH71966 standard; cDNA; 470 BP.

XX AAH71966;
 AC
 XX 19-SEP-2001 (first entry)
 DT
 XX

DE Human cervical cancer marker nucleic acid 3240.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

XX WO200142467-A2.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000WO-US033312.

XX 08-DEC-1999; 99US-0168681P.

XX 21-DEC-1999; 99US-0171350P.

XX 14-MAR-2000; 2000US-0189315P.

XX 09-JUN-2000; 2000US-0203791P.

XX 21-JUL-2000; 2000US-0220114P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Deede J, Berger A, Zhao X;

DR WPI, 2001-375006/39.

PT New isolated nucleic acid for diagnosing and treating cervical cancer and
 for assessing and detecting compounds for treating the cancer.

PS Claim 1; Page 636; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy
 CC

SQ Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

Query Match 33.9%; Score 426.6; DB 4; Length 470;
 Best Local Similarity 98.7%; Pred. No. 2.4e-116;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 3159.5 Seconds
(without alignments)
4832.489 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSGCAGMBPFRFLFARG.....VATFYPPLSAVKRGYLTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DRV=xlp
-Q/cgn2_1/USPTO.epool_p/US10798532/runat_29112004_112453_15901/apd_query.faeta_1.1166
-DB=EST -QFMT=fastap -SUFPLX=rat -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITs=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532.cgcen_1_1_6628@runat_29112004_112453_15901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -XGAPOP=6
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:.*
1: gb_estc1:.*
2: gb_estc2:.*
3: gb_htc:.*
4: gb_estc3:.*
5: gb_estc4:.*
6: gb_estc5:.*
7: gb_estc6:.*
8: gb_gsa1:.*
9: gb_gsa2:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 2171 | 98.4 | 1763 | 3 | AK013347 Mus muscu |
| 2 | 1711 | 77.6 | 2846 | 3 | AK030152 Mus muscu |
| 3 | 1521.5 | 69.0 | 1079 | 4 | BM468107 AGENCOURT |
| 4 | 1271.5 | 57.6 | 1015 | 6 | BY712344 BY712344 |
| 5 | 1242 | 56.3 | 917 | 4 | BM451184 AGENCOURT |
| 6 | 1169 | 53.0 | 823 | 7 | C0771206 testis ES |
| 7 | 1162 | 52.7 | 779 | 5 | BU447831 BU447831 |
| 8 | 1158 | 52.5 | 896 | 7 | C0773091 testis ES |
| 9 | 1133 | 51.4 | 781 | 7 | CN232453 CN232453 |

| | | | | | | |
|----|--------|------|------|---|----------|--------------------|
| 10 | 1107 | 50.2 | 932 | 5 | BU541306 | BU541306 AGENCOURT |
| 11 | 1106 | 50.1 | 717 | 5 | BU385761 | BU385761 603582650 |
| 12 | 1091 | 49.5 | 899 | 2 | BE895119 | BE895119 601436041 |
| 13 | 1079 | 48.9 | 772 | 4 | BI562420 | BI562420 60326316 |
| 14 | 1075 | 48.7 | 794 | 5 | BU205357 | BU205357 603104559 |
| 15 | 1070.5 | 48.5 | 834 | 7 | CF593363 | CF593363 AGENCOURT |
| 16 | 1068.5 | 48.4 | 875 | 4 | BI259568 | BI259568 602968330 |
| 17 | 1063 | 48.2 | 888 | 5 | BQ225071 | BQ225071 AGENCOURT |
| 18 | 1050 | 47.6 | 657 | 4 | BG547675 | BG547675 602575520 |
| 19 | 1050 | 47.6 | 682 | 4 | BI333581 | BI333581 602977069 |
| 20 | 1025 | 46.5 | 844 | 4 | BG474616 | BG474616 602517423 |
| 21 | 1014 | 46.0 | 602 | 7 | CO737403 | CO737403 SLT02c17 |
| 22 | 1005.5 | 45.6 | 876 | 5 | BU184346 | BU184346 AGENCOURT |
| 23 | 990.5 | 44.9 | 906 | 4 | BG386905 | BG386905 602454704 |
| 24 | 968.5 | 43.9 | 966 | 5 | BU439278 | BU439278 604147528 |
| 25 | 967 | 43.8 | 631 | 5 | BU112717 | BU112717 603131421 |
| 26 | 966.5 | 43.8 | 1003 | 5 | BQ222020 | BQ222020 AGENCOURT |
| 27 | 962 | 43.6 | 755 | 5 | BU371858 | BU371858 603568233 |
| 28 | 955 | 43.3 | 699 | 5 | BU070284 | BU070284 UI-M-FRO- |
| 29 | 944 | 42.8 | 604 | 5 | BX261107 | BX261107 BX261107 |
| 30 | 909.5 | 41.2 | 716 | 7 | CF745774 | CF745774 UI-M-GVO- |
| 31 | 900 | 40.8 | 570 | 5 | BU293329 | BU293329 604166812 |
| 32 | 893 | 40.5 | 670 | 7 | CN359673 | CN359673 170004245 |
| 33 | 889 | 40.3 | 928 | 5 | BQ424438 | BQ424438 AGENCOURT |
| 34 | 880.5 | 39.9 | 824 | 2 | BE568528 | BE568528 601342425 |
| 35 | 870.5 | 39.5 | 820 | 7 | CK597557 | CK597557 AGENCOURT |
| 36 | 862 | 39.1 | 672 | 5 | BX261106 | BX261106 BX261106 |
| 37 | 858.5 | 38.9 | 1663 | 6 | CD510246 | CD510246 CDAL00-G0 |
| 38 | 855 | 38.8 | 785 | 6 | CB990192 | CB990192 AGENCOURT |
| 39 | 850 | 38.5 | 523 | 2 | BF299410 | BF299410 IL2-MT020 |
| 40 | 828.5 | 37.6 | 1063 | 5 | BU356643 | BU356643 603474124 |
| 41 | 818.5 | 37.1 | 743 | 4 | BI562440 | BI562440 603256324 |
| 42 | 815 | 36.9 | 763 | 7 | CK364431 | CK364431 AGENCOURT |
| 43 | 798 | 36.2 | 933 | 2 | BE886040 | BE886040 601507192 |
| 44 | 796 | 36.1 | 986 | 5 | BU751496 | BU751496 CH4#002_E |
| 45 | 793 | 35.9 | 484 | 5 | BX281211 | BX281211 BX281211 |

ALIGNMENTS

RESULT 1
AK013347
LOCUS
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:281045604 product:kinase interacting with
leukemia-associated gene (bractinin), full insert sequence.

ACCESSION AK013347
VERSION AK013347.1 GI:12850651
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 11042159
PUBMED 2049374

REFERENCE
AUTHORS Shizata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Komno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiwa, K.,
Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M.,

| | | | |
|------------|------|---|--|
| Db | 939 | ATATTGGCAAGTAAGCAGTGGTGAATCCGCAATTCACGCTTACCTCAGAGACTT | 998 |
| Qy | 281 | IlleySerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys | 300 |
| Db | 999 | ATCAAAAGCATGCTTCATGATGACCCGGGAGAAAGATCCCTGCTGAGATGGCATGTGTC | 1058 |
| Qy | 301 | SeiProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr | 320 |
| Db | 1059 | AGCCCATTCCTTAGCATTCCTTTTTCGCCCTCATATTTGAAGATCTGGTATGCTCCGACT | 1118 |
| Qy | 321 | ProValLeuArgLeuLeuAsnValLeuAspAspAspPheTyrLeuGluYasnGluGluGly | 340 |
| Db | 1119 | CCAGGCTCAGGCTCTCTCAATGCTGGATGTGATTCCTTGAAAGAGATGAATAT | 1178 |
| Qy | 341 | GluAspValValGluAspValValLeuGluGluCysGlnIleTyrGlyProValValSerLeu | 360 |
| Db | 1179 | GAAAGTGTGTGTAAGATGCTGAAGAGAGAGTCTCGAAATATGACCAAGTGGTTCTCTG | 1238 |
| Qy | 361 | LeuValProValGluAsnProGlyValArgGlyGluValPheValGluTyrAlaAsnAlaGly | 380 |
| Db | 1239 | CTGTGTCCAAAGGAAAAATCTCTGGCAGAGAGACAAAGTCTTGATGATACCGCAACCTGGT | 1298 |
| Qy | 381 | AspSerIysAlaIleAlaGlnIleLeuLeuThrGlyArgMetPheAspGlyIlePheValVal | 400 |
| Db | 1299 | GATTCCTCAAGCTGCTCAGAAAGTGTGTGCTGGAGAGAAATTTGACAGGAAATTTTGTGTG | 1358 |
| Qy | 401 | AlaThrPheTyrProLeuSerAlaTyrIleYsaArgGlyTyrLeuTyrGlnThrLeuLeu | 419 |
| Db | 1359 | GCTACATTCCTACCCGCTAGTGCCTCCACAAAGGGGATATCTTTATCAAACTTGGTT | 1415 |
| RESULT 2 | | | |
| AK030152 | | 2846 bp | mRNA |
| LOCUS | | | linear |
| DEFINITION | | | HTC 03-APR-2004 |
| | | | Mus musculus adult male testis cDNA, RIKEN full-length enriched |
| | | | library, clone:493340107 product:kinase interacting with |
| | | | lektema-associated gene (stachmin), full insert sequence. |
| ACCESSION | | | AK030152 |
| VERSION | | | AK030152.1 GI:26326136 |
| KEYWORDS | | | HTC; CAP trapper. |
| SOURCE | | | Mus musculus (house mouse) |
| ORGANISM | | | Mus musculus |
| | | | Eumaliaota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| | | | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| MEDLINE | | | |
| PUBMED | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL</ | | | |

| | |
|---------------------|--|
| TITLE | PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) |
| JOURNAL | |
| REFERENCE | |
| AUTHORS | 5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) |
| TITLE | |
| JOURNAL | |
| REFERENCE | 6 (bases 1 to 2846) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hagiwara, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohashi, Y., Ozaki, Y., Saito, R., Satoh, K., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Sehiho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp), URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) |
| COMMENT | cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers |
| FEATURES | |
| source | 1. .2846 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM_DB:4933401D07" /db_xref="taxon:10090" /clone="4933401D07" /sex="male" /tissue_type="testis" /clone_id="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 238. .1230 /note="unannotated protein product; kinase interacting with leukemia-associated gene (stathmin) (MGI:1341908) putative" /codon_start=1 /protein_id="BAC26812.1" /db_xref="GI:26326137" /translation="MTLTGFTTFHSNPVSRCLLELDVSVSELLYSHQCSMMW MIOHGRAPBALAFHLEHGVYHVDLPRNIIAMSEAGCILLDFPGSGNDGVKK IQDNGYARPEALONCLAQGLSDPTCSAYVILMSGLITLDFPSPMKIKHYRSGE WKANSSAIIDHIFASKAVVNAAPAHYLRILISMLEDHDDGRRIPAEMALCSPPFSIF FAPRIDEVLWLPVRLNLNVLDLDDYLENDDEYEDVDEKCEQKGPVAVSLVPE NPGGGVFEVYANAGDSKAQKILTGMPDGKFEVATVPPLSAVKRGYLVTL" 2823. .2828 /note="putative" 2846 /note="putative" |
| ORIGIN | |
| Alignment Scores: | |
| Prod. No.: | 1.87e-167 |
| Score: | 1711.00 |
| Percent Similarity: | 99.3% |
| | Length: 2846 |
| | Matches: 326 |
| | Conservative: 2 |

Best Local Similarity: 98.79% Mismatches: 2
 Query Match: 77.56% Indels: 0
 DB: 3 Gaps: 0

US-10-798-532-4 (1-419) x AK030152 (1-2846)

```

QY 90 ValThrLeuTyGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
DB 238 ATGACTTTATATGAGATCTTTACCATACACTTCTCCCAATGTGCACACCGCTCTCG 297
QY 110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuLeuLeuLeuLeuLeu 129
DB 298 CTGCTTGAACCTCTGGATGTCAGTGTTCGAAATCTTTATATTCCTCATTCAGGCG 357
QY 130 CysSerMetTrpMetIleGlnHisCysValArgAspValLeuGluValLeuValPheLeu 149
DB 358 TGGTCCATGTGATGATACAGCATCTGTGCCGAGATGTTCTGGAGGCCCTCGTTTCTT 417
QY 150 HisHisGluGlyTyValHisAlaAspLeuLeuProArgAsnIleLeuTrpSerAlaGlu 169
DB 418 CACCATGAGGGGTACCTCATGCAGACCTCAAAACGAAACATCTCTGGAGTGGGAG 477
QY 170 AsnGluCysPheLeuLeuLeuAspPheGlyLeuSerPheLeuGluValAsnGluAspVal 189
DB 478 AATGAGTGTTTAAGCTTATGACTTTGACCTCAGCTTCAAGACGCAATCAGACGTA 537
QY 190 LysTyrlleGlnThrAspGlyTyValArgAlaProGluValGluLeuGlnAsnCysLeuAla 209
DB 538 AGTATATTCAGACGACGCGTATCGCGCTCCGAAAGCAGAGCTGCAGAACTGCTGGCC 597
QY 210 GlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGly 229
DB 598 CAGGCGGCGCTGAGAGTATCAGATGATACCTCAGCTGTGATCTGTGAGACCTCGGA 657
QY 230 IleIleLeuLeuGluLeuLeuPheSerGlyMetLeuLeuLeuHisPheValArgSerGlnGlu 249
DB 658 ATCATTTTCTGAAATGTTCTCGAAGATGAACTTAACATACATACAGTTCAGGAA 717
QY 250 TrpValAsnSerSerAlaIleIleAspHisIlePheHisAspLeuValValAsn 269
DB 718 TGGAGGCAAAACGTTCTGTATATTTGATCATATATTTGCTCCGTAAGAGTGGTGAAT 777
QY 270 AlaAlaIleProAlaTyHisIleLeuArgAspLeuIleLysSerMetLeuHisAspAspPro 289
DB 778 GCGCAATTCACAGCCTATCCTCAGAGACCTTATCAAAAGCATGCTTATGATGACCG 837
QY 290 SerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAla 309
DB 838 GCGAAGAAATCCCTCTGATGATGCAATGTGCAAGCCCATTTCTTAGCATTCCTTTTGC 897
QY 310 ProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeuValLeu 329
DB 898 CCTCATATGAGATCTGGTGAATGCTTCGACCTCCAGTGTCTCAGGCTCCCATATGTCGTG 957
QY 330 AspAspAspTyrlleuGluAsnGluGluTyrlleuAspValValGluAspValLysGlu 349
DB 958 GATGATGATTAACCTTCAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAG 1017
QY 350 GlnCysGluLysTyrlleuTyrlleuProValValSerLeuLeuValProLysGluAsnProGlyArg 369
DB 1018 GACTGTCAAGAAATATGACCAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1077
QY 370 GlyGlnValPheValGluTyrlleuAlaAsnAlaGlyAspSerLysAlaIleGlnLysLeu 389
DB 1078 GGAACAATCTTCTGATGATCGGAAACGCTGATGATTCCAAACCTGCTCAGAACTTGTG 1137
QY 390 ThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyrlleuLeuSerAlaTyrlleu 409
DB 1138 ACTGGAGAGATGTTTGAACGGAATGTTGTGTGCTTACATTTTCAACCGCTGATGCTTAC 1197
QY 410 LysArgGlyTyrlleuTyrlleuGlnThrLeuLeu 419
DB 1198 AAGAGGGGATATCTTATCAAACTTGTCTT 1227

```

RESULT 3
 BM468107 1079 bp mRNA linear EST 05-FEB-2002
 LOCUS 5' mRNA sequence.
 DEFINITION AGENCOURT 6432293 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535653
 ACCESSION BM468107 GI:18517149
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
 1 (bases 1 to 1079)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strusberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
 Plate: LLM12224 row: b column: 06
 High quality sequence stop: 754.
 Location/Qualifiers
 1..1079
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5535653"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb."

FEATURES

source
 1..1079
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5535653"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb."

ORIGIN

Alignment Scores:
 Pred. No.: 3,05e-148 Length: 1079
 Score: 1521.50 Matches: 309
 Percent Similarity: 92.86% Conservative: 3
 Best Local Similarity: 91.96% Mismatches: 17
 Query Match: 68.97% Indels: 7
 DB: 4 Gaps: 3

US-10-798-532-4 (1-419) x BM468107 (1-1079)

```

QY 90 ValThrLeuTyGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeu 109
DB 2 GATGACTTTATATGAGATCTTTACCATACACTTCTCCCAATGTGCACACCGCTCTCG 61
QY 110 LeuLeuGluLeuLeuAspValSerValSerGluLeuLeuLeuLeuLeuLeuLeuLeu 129
DB 62 TTGCTTGAACCTCTGGATGTCAGTGTTCGAAATCTTTATATTCAGTCCAGCAGGCT 121
QY 130 CysSerMetTrpMetIleGlnHisCysValArgAspValLeuGluValLeuValPheLeu 149
DB 122 TGGTCCATGTGATGATACAGCATCTGTGCCGAGATGTTTGAAGGCCCTTCTTTCTT 181
QY 150 HisHisGluGlyTyValHisAlaAspLeuLeuProArgAsnIleLeuTrpSerAlaGlu 169
DB 182 CATCATGAGGCTATGTCTCATCGGACCTCAAAACGTAACATATTTGGATGTCAGAG 241
QY 170 AsnGluCysPheLeuLeuLeuAspPheGlyLeuSerPheLysGluGlyValAsnGluAspVal 189
DB 242 AATGATGTTTAAACTCATGACTTGTGACTTGAATCAAGAAAGCAATCAGGATGTA 301
QY 190 LysTyrlleGlnThrAspGlyTyValArgAlaProGluValGluLeuGlnAsnCysLeuAla 209

```

```

Db      302 AAGTATATTCAGACACACGGGTTCCGGCTCCAGAAAGCAAAATTCGAGGCTGGCC 361
Qy      210 G|A|A|G|L|V|E|N|G|I|N|S|E|A|P|T|H|U|C|Y|E|T|H|S|E|A|V|A|L|A|S|P|E|U|T|P|S|E|U|G|Y 229
Db      362 CAGGCTGGCCCTGAGAGTGTATACAGATGTAACCTCAGCTGTATATCTGTGGACCTTAGA 421
Qy      230 I|L|L|L|E|U|L|E|U|G|I|U|N|E|T|P|H|E|S|E|G|Y|M|E|L|Y|E|U|L|Y|H|I|S|H|Y|A|L|A|S|E|S|E|G|I|N|G|U 249
Db      422 ATCATTTTATTCGAAATGTTTCTCAGCAATGAACTGAAACATTCAGTCAGATTCCTCAGAA 481
Qy      250 T|P|L|Y|A|L|A|A|N|S|E|S|E|S|A|L|I|L|L|E|A|P|H|I|L|E|P|H|E|A|S|E|S|I|Y|S|A|V|A|L|A|A|N 269
Db      482 TGGAGAGCAAAACAGTCTCTGCTATTAATGATCATATTTTCCAGTAAAGCAGGCGTAAT 541
Qy      270 A|A|A|L|L|P|P|A|L|L|P|P|A|L|L|Y|H|I|S|E|A|A|A|P|L|L|L|E|Y|S|E|T|N|E|U|H|I|S|A|P|P|P|O 289
Db      542 GCGCAATTCACACCTATCACCCTATCCTATCCTATCCTATCCTATCCTATCCTATCCTA 601
Qy      290 S|E|A|A|G|A|G|L|E|P|A|L|G|U|E|T|A|L|E|U|C|Y|S|E|S|P|H|E|S|E|S|I|L|E|P|H|E|A|L 309
Db      602 AGCAGAAAGATTCCTCTCTGAAATGGCATTTGCGACCCATCTTTAGCATTCCTTTTGGC 661
Qy      310 P|H|I|H|I|L|E|G|U|A|P|L|E|U|V|A|L|N|E|U|P|H|O|T|H|P|R|O|V|A|L|E|U|A|G|L|E|U|A|N|V|A|L|E|U 329
Db      662 CCTCATATTGAAGATCTGCTCATGCTTCCACATCCAGTCTAAGACTGCTGAATGTGCTG 721
Qy      330 A|A|P|A|A|P|A|P|Y|L|E|U|G|Y|A|N|G|U|G|U|G|U|T|Y|G|L|U|A|P|V|A|L|G|U|A|P|V|A|L|Y|G|U 349
Db      722 G|A|T|A|G|A|T|T|Y|T|T|G|A|A|T|G|A|A|G|A|T|T|G|A|A|G|A|T|T|G|A|A|G|A|T|T|G|A|A|G|A 781
Qy      350 G|U|C|Y|G|I|N|Y|T|Y|G|Y|P|R|O|V|A|L|S|E|L|E|U|V|A|L|P|R|O|Y|G|U|A|N|P|R|O|Y|G|Y|A|G 369
Db      782 G|A|G|T|C|A|A|A|A|A|T|G|A|C|C|A|G|G|G|T|A|T|C|T|G|T|C|T|G|T|C|T|G|T|C|T|G|T|C|T|G|G|C|A 841
Qy      370 G|Y|G|I|N|V|A|L|P|H|E|V|A|L|U|T|Y|A|L|A|A|N|A|L|G|Y|A|P|S|E|L|Y|S|A|V|A|L|G|I|N|Y|L|E|U 388
Db      842 G|G|C|A|G|A|G|T|C|T|T|G|N|G|A|G|T|A|T|G|C|A|A|T|G|T|G|N|G|T|A|T|C|A|A|A|G|T|G|C|G|C|A|A|A|T|T|A 901
Qy      389 L|E|U|T|H|G|Y|A|G|---|M|E|T|P|H|E|A|P|G|Y|L|Y|S|P|H|E|V|A|L|A|L|A|---|T|H|P|H|E|Y|P|R|O|L|E|U|--- 406
Db      902 C|T|G|A|C|G|G|A|A|N|A|N|G|A|T|G|T|T|G|A|T|G|G|G|A|A|T|T|G|N|G|T|G|N|G|C|C|T|A|C|A|T|C|C|C|C|C|T|G 961
Qy      407 -S|E|A|L|A|T|Y|L|Y|E|A|G|Y|---|T|Y|L|E|U|Y|Y|G|I|N|T|H|L|E|U 418
Db      962 G|A|G|T|G|C|T|T|A|C|A|G|A|G|G|G|A|T|T|T|C|T|G|T|T|C|A|A|C|T|T 1001

RESULT 4
BY12344 1015 bp mRNA linear EST 17-DEC-2002
LOCUS BY12344
DEFINITION Mus musculus cDNA clone 2810454004 5', mRNA sequence.
ACCESSION BY12344
VERSION BY12344.1 GI:27123621
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
1 (bases 1 to 1015)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamakawa, Y., Nogi, A.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogi, A.,
Schonbach, C., Gojodori, T., Balderelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Choitha, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

```

Numata, K., Okido, T., Pavan, M.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shmida, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, T.,
Wells, C., Wilming, L.G., Wyszewski, B., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kitahara, T., Kono, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y.,
Itoh, M., Kagawa, T., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL MEDLINE PUBMED

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,
Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,
Imocani, K., Ishii, Y., Itoh, M., Kagawa, T., Kawai, J., Kojima, Y.,
Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,
Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,
Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES

source

Location/Qualifiers
1..1015
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/cclone="2810454004"
/tissue_type="whole body"
/dev_stage="10, 11 days embryo"
/clone_lib="RIKEN full-length enriched, 10, 11 days embryo
whole body"

ORIGIN

Alignment Scores:

Pred. No.: 4,246-122 Length: 1015
Score: 1271.50 Matches: 251
Percent Similarity: 89.51% Conservative: 5
Best Local Similarity: 87.76% Mismatches: 23
Query Match: 57.64% Indels: 7
DB: 6 Gaps: 4

US-10-798-532-4 (1-419) x BY712344 (1-1015)

QY 1 MetAlaGlySerGlyCysAlaATPGLyAlaGluProProArpPheLeuGluAlaPheGly 20
 Db 159 ATGGCGGGGCTCCGGCTGCGGGGGGCGGAGCCGCCGCTTCTTGAGGCGCTTCGGG 218

QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 Db 219 CGGCTGTGGAGGCTCCAGAGCCGCTGGGAGCGGCTCGTGGGCTCGGCTGTCGGG 278

QY 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
 Db 279 CGGTGCTGGGTACCCCGGGCTGCGCCCGCGGCGCTCAAGCATTCCTGCTCGGGA 338

QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgPheGluArgAlaAlaLeu 80
 Db 339 ACCACCGGGGCTCGGGCTCGCGCGGAGTATGTTTCCGCAAGAGAGCGGCGCTG 398

QY 81 GluGlnLeuGlnGlyH1sArgAsnIleValThrLeuTyrGlyValPheThrIleH1sPhe 100
 Db 399 GANCACTTGACAGGCTCAGAGAACATCGTACTTTATATGAGTCTTTACATACACTTC 458

QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
 Db 459 TCTCCCAATGTGCATCACGCTGTCTGCTGCTGAACCTCTGATGTCAGTGTTCGGA 518

QY 121 LeuLeuLeuTyrSerSerH1sGlnGlyCysSerMetTrpMetIleGlnH1sCysAlaArg 140
 Db 519 TTGCTTTATATTCAGTCACTACGGGCTGCTCATGTGATGATGATACGACTGTGCCA 578

QY 141 AspValLeuGlnAlaLeuAlaPheLeuH1sH1sGlnGlyTyrValH1sAlaAspLeuLys 160
 Db 579 GATGTTCTGGAGGCGCTCGCTTTCTTACCATGAGGGCTACGCTCATGACAGCTCAAA 638

QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
 Db 639 CCACGAACATCTGTGAGAGTGGGAGAAATGATGTTTAAAGTTTATGACTTGGACTC 698

QY 181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
 Db 699 AGCTTCAAGAGAGGCAATCAGAGACGTAAATATTCAGACACAGCGGATTCGGCTCCT 758

QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 Db 759 GAAGCAGAGCTGCAGAACTGCTTGCCCGAGCGCGCTGCAGAGTATACAGAAATGATCC 818

QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
 Db 819 TCAGCTGTGATGTGCGGGCTCGGAATCACTTACTGAGATGTTCTCAGGAATGAAG 878

QY 241 LeuLys---H1sThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
 Db 879 CTGAAATATCATCATCATCAGCA-----TGGAGAGCA---AGCACTCGCATCATTCG 929

QY 260 H1sIlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrH1sLeuArgAsp 279
 Db 930 CATATATTTT-----GCAGTGAGCAGATGTGTTCGCAATTCAGCCATCACTCAGAC 980

QY 280 LeuIleLysSerMetLeu 285
 Db 981 GCTTATCAGAGCATGCTC 998

RESULT 5
 LOCUS BM451184 917 bp mRNA linear EST 05-FEB-2002
 DEFINITION AGENCOURT_6392719 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5495159
 5', mRNA sequence.
 ACCESSION BM451184
 VERSION BM451184.1 GI:18500224
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1M12121 row: 3 column: 24
 High quality sequence stop: 718.

FEATURES
 source
 1..917
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5495159"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_67"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN
 Alignment Scores:
 Pred. No.: 4,47e-119 Length: 917
 Score: 1242.00 Matches: 240
 Percent Similarity: 99.59% Conservative: 0
 Best Local Similarity: 99.59% Mismatches: 1
 Query Match: 56.30% Gaps: 0
 Indels: 0

US-10-798-532-4 (1-419) x BM451184 (1-917)

QY 179 GlyLeuSerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArg 198
 Db 22 GGGCTTAAGCTTCAAGAGAGCAATCAGATGTAAATATTCAGACACAGCGGTATCGG 81

QY 199 AlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlu 218
 Db 82 GCTCCAGAGCAAGATTGCAAAATTCCTTGCGCCAGCGCTGCAGAGTATACAGAA 141

QY 219 CysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGly 238
 Db 142 TGTACTCAGCTGTGATGTGTGAGCGCTAGGAATCATTTTACTGGAATGTTCTCAGGA 201

QY 239 MetLysLeuLysH1sThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIle 258
 Db 202 ATGAACCTGAACATACATCAGATCTCAGAGATGAAGCAAGCAAGTTCGTATATTT 261

QY 259 AspH1sIlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrH1sLeuArg 278
 Db 262 GATCACAATTTTCCAGTAAACAGATGTGATGCCAGCAATTCAGCCATACCTAAG 321

QY 279 AspLeuIleLysSerMetLeuH1sAspAspProSerArgArgIleProAlaGluMetAla 298
 Db 322 GACCTTATCAAAAGCATGCTCATGATGATCAAGCAGAAAGATTCGTGTAATGCGA 381

QY 299 LeuCysSerProPhePheSerIleLeuProAlaProH1sIleGluAspLeuValMetLeu 318
 Db 382 TTGTGAGCCCATTTCTTAAGATTCCTTTTGCCCTCATATTTGAAGATCTGTGATGCTT 441

QY 319 ProThrProValLeuArgLeuLeuAsnValIleAspAspAspTyrLeuGlyAsnGluGlu 338
 Db 442 CCACCTCCAGTCTAAGACTGCTGAATGTCTGAGATGATGATTAATCTTGAGAAATGAAG 501

QY 339 GIUTYRGUASPVALVALGLUASPVALLYSGUUGLUCYSGUULYSTRYRGYPROVALVAL 358
DB 502 GATATAGAGAGATTGTAGAGATGTAAAGAGAGATGTCAAAAATATATGACAGCGTA 561
QY 359 SerLeuLeuValProLySGUASPVALLYSGUUGLUCYSGUULYSTRYRGYPROVALVAL 378
DB 562 TCTCACTTGTTCCTCAAGAGAAATCTCGCAGAGCAAGTCTTTTCAATATGCAAAAT 621
QY 379 AAGAGASPLeuValValAGUULYSGUUGLUCYSGUULYSTRYRGYPROVALVAL 398
DB 622 GCTGGAGATTCAG 681
QY 399 Val 418
DB 682 GTTGTGCTCAACATTCACCGCTGAGTGTGCTCAACAGAGAGAGAGAGAGAGAGAGAG 741
QY 419 Leu 419
DB 742 CTT 744

RESULT 6
LOCUS CO771206 823 bp mRNA linear EST 04-AUG-2004
DEFINITION testis EST02303 Testis cDNA library Gallus gallus CDNA 3', mRNA
ACCESSION CO771206
VERSION CO771206
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 823)

REFERENCE
AUTHORS Shin,J.
TITLE Testis cDNA library
JOURNAL Unpublished (2004)
COMMENT Contact: JiHee Shin
Division of Animal Genetic Engineering
School of Agricultural Biotechnology, Seoul National University
San 56-1, Sillim-dong, Gwanak-gu, Seoul, 151-742, Korea
Tel: +82-31-479-7557
Fax: +82-31-479-7550
Email: paecllo@hanmail.net
Seq primer: T3

FEATURES

source
1. .823
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Korean Native chicken"
/db_xref="taxon:9031"
/sex="Male"
/tissue_type="Testis"
/dev_stage="26weeks"
/clone_lib="Testis cDNA library"
/note="Vector: Uni-ZAP XR vector; Site_1: XhoI; Site_2:
ECORI. The libraries were synthesized using the
ZAP(R)-cDNA synthesis method (Stratagene). cDNA was
prepared, size-fractionated and inserted into the Uni-ZAP
XR vector using XhoI linker-primer and EORI adaptor.
After in vivo excision with E. coli strain SOLR, cDNA
libraries from testis contain each insert size range of
0.5-3kb (n=20), and 0.5-3.5kb (n=20)."

ORIGIN

Alignment Scores:
Pred. No.: 1.68e-111 Length: 823
Score: 1169.00 Matches: 229
Percent Similarity: 84.38% Conservative: 14
Best Local Similarity: 79.51% Mismatches: 23
Query Match: 52.99% Indels: 22
DB: 7 Gaps: 2

US-10-798-532-4 (1-419) x CO771206 (1-823)
QY 47 GlycerProPro-----GLYAlaLeuArgInPheLeuProProGlyThrThrGly 63
DB 6 GAGAGCTCCAGCCGGGTGGCCGCGCTTAAAGATAGTGAATCCCCGGGGCTGCAGGAAT 65
QY 64 AlaAlaAlaSerAlaAlaGlyUtyrGlyPheArgLySGUASPVALVALAGUUGLULeu 83
DB 66 TCGGACAGAGGC----- 77
QY 84 GInGlyHISArgAsnIleValThrLeuTyrglyValPheThrIleHisPheSerProAsn 103
DB 78 -----CGCAACATGTGATCTGTACCGCCGGTTCACCAACCATCTACCGCCAAC 128
QY 104 ValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGluLeuLeu 123
DB 129 GGCCTGCTCCCGCTGCTGCTGCTGAGCTGTGAGATACAGCGTGTGAGCTGCTGCTG 188
QY 124 TyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAspValLeu 143
DB 189 CACTCCAGCAACCAAGGCTGCTCCATGTGATGATCCAGCACTGCGCGGAGCTGCTG 248
QY 144 GUAAlaLeuAlaPheLeuHisGlyGlyTyrglyValHisAlaAspLeuLeuProArgAsn 163
DB 249 GAAGCCCTGGCCTTCTGCTGCACCAAGGCTACGAGCAAGCACTGAAAGCCAGCAAC 308
QY 164 IleLeuTrpSerAlaGluAsnGluCysPheLeuLeuAspPheGlyLeuSerPheLeu 183
DB 309 ATCTGTGTGAGGCGGAG 368
QY 184 GIUGLYAsnGlnAspValLystrYrIleGlnThrAspGlyTyrglyAlaProGluAlaGlu 203
DB 369 GAGGGAGATCAGAGATGTGAATATATTTCAACAGCGGATTCGGGCTCCAGAGCAGAA 428
QY 204 IeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSerAlaVal 223
DB 429 CTGCAGAACTGCTGCTGACAGAGAGGCTCCAGAGTGAAGAGATGATCTGCTG 488
QY 224 AspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLeuLeuHis 243
DB 489 GATCTGTGTGAGCTGTGAATGTGTTTACTGTGAAGAGTTCTCAGATATAACTGAACAT 548
QY 244 ThrValArgSerGlnGlyTrpLyAlaAsnSerSerAlaIleIleAspHisIlePheAla 263
DB 549 ACGTCCCAATCTCAGAGATGGAAGCAACAGTTTGCATCATGATCGCATTTTTC 608
QY 264 SerLyAlaAlaValAlaAsnAlaIleProAlaTyrglyHisLeuArgAspLeuIleYSer 283
DB 609 ACTGAAGGGGTGTAAATTCAGCCATTCAGCTTATCAGCTCAGAGACCTTATTAAGC 668
QY 284 MetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhe 303
DB 669 ATGCTTCAATGTGACCAAGCAGAGCTCTGTGAAAGGCTTATGAGCCCATTC 728
QY 304 PheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeu 323
DB 729 TTCAGCATTCCTTTGCTCCCATATTTGAAGATTTGGTGAATCTCCACAGCTGTGCTG 788
QY 324 ArgLeuLeuAsnValLeuAspAsp 331
DB 789 AGCTGTGCTGAATGTTCTTAACGAT 812
RESULT 7
LOCUS BU447831 779 bp mRNA linear EST 29-NOV-2002
DEFINITION 603767571F1 CSQRBN13 Gallus gallus CDNA clone ChEST695n17 5', mRNA
ACCESSION BU447831
VERSION BU447831.1 GI:25937142
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 779)
 Boardman, P. E., Sanz-Eguero, J., Overton, I. M., Burt, D. W., Boech, E.,
 Fong, W. T., Tickle, C., Brown, W. R. A., Wilson, S. A. and Hubbard, S. J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 JOURNAL MEDLINE 2235534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..779
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer"
 /db_xref="taxon:9031"
 /clone="CHEST695n17"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_1lb="CSEORBN13"
 /note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
 EcoRI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunt-ended, ligated to NotI adapters, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN
 Alignment Scores:
 Pred. No.: 8,36e-111 Length: 779
 Score: 1162.00 Matches: 224
 Percent Similarity: 93.02% Conservative: 16
 Best Local Similarity: 86.82% Mismatches: 18
 Query Match: 52.67% Indels: 1
 Gaps: 0
 DB: 5

US-10-798-532-4 (1-419) x BU447831 (1-779)
 QY 158 Aspleuylsefrcarxanilleleutpseralaglunsglucyspheyseuileap 177
 Db 6 GACCTGAAGCAGCAACATCTCTGTGAGCGCGAGAGAGAGCTTTAAGCTCATTTGAC 65
 QY 178 Phegllyleuserhellysgluyanglnasvallyyrtllleqlnthrapgllytyr 197
 Db 66 TTTGGACTTAGCTTCAAGAGGGGAATCAGGAAGTGAATATTTCAACACAGCGGGTAT 125
 QY 198 ArgAlProgluagluleuglnaenCysleuAlaglulaglyleuglnaserApTtr 217
 Db 126 CGGGCTCCAGAGGCAAGACTGCGAACTGCTTGCACAGCAGGAGGCTCCAGATGAGAGC 185
 QY 218 GluCyThserAlaValaAspleuITPserleuglylleleleuengluwetPheser 237
 Db 186 GAATGACCTCTGCTGTGATCTGTGAGCTCTGGAATTTTAACTGGAAATGTTCTCA 245
 QY 238 GlyMetlyleuylshlstrValaArgserGlnGluTTPlyAlaenSerSerAlaIle 257

Db 246 GGATGAAATCGAATACATCACTCCATCTCAGGAATGAGACAAACAGTTCCCATC 305
 QY 258 IleaphsiilePheAlaserlysaIaValaAenAlaIleProAlaTyrihsleu 277
 Db 306 ATCGATCGATTTTGGCCAGTGAAGGGGTGTTATTCAGCATTCAGGCTTATCATCCTC 365
 QY 278 ArgAspleuilelysermetleuHlsapAspProserArgArgIleProAlaIleuMet 297
 Db 366 AGAGACTTATTAAGCAAGCTCTCATTCGTGACCAAGCAGCGAGCTGCTGTAAG 425
 QY 298 AlaLeuCyserProPhePheSerIleProPheAlaProHisIleGluAspleuValMet 317
 Db 426 GCTTATGAGCCCATCTTCTTCAGCATTCCTTGTGCTCCCATATGAAGATTCGGTGATG 485
 QY 318 LeuProtnProValleuArgleuLeuAsnValleuAspAspTyrlenglyasnglu 337
 Db 486 CTCGCCAGCGCTGTGCTGAGCTGCTGAATGTTCTTAAGCAGATTCCTCTGACAGTGAA 545
 QY 338 GluGluTyrgluAspValaIgluAspVallysgluyanglnucyeglnlytyrglyProVal 357
 Db 546 GAAGAAATCAGAAATATCCTCGAAGACATTAAGGAGAGGTGCAGAAATATGACCGGTG 605
 QY 358 ValserleuLeuValProlysgluAsnProgluArggluGlnValPheValGluTyrla 377
 Db 606 GTTTCCTGCTTATTCAGAGAAATCTGTAAAGGCCAAAGCTTTGT-GAATATGCA 664
 QY 378 AsnAlaGlyAspSerlysaIaIaGlnlyseuLeuThrGlyArgMetPheAspGlylys 397
 Db 665 AATGCTGTGATTCGAAGAGCTGCCAAGAAATGCTGACGTGGAAGATTTTGATGACGAAG 724
 QY 398 PheValaIaIaThPheTyProleuserAlaTyrlaGlylytyrleuTy 415
 Db 725 TCTGTCGTGCTACGTTTACCACCTGAGTGTCTTAAGAGAGATATCTGTAC 778

RESULT 8
 LOCUS CO773091 896 bp mRNA linear EST 04-AUG-2004
 DEFINITION testis cDNA library Gallus gallus cDNA 3', mRNA
 ACCESSION CO773091
 VERSION CO773091.1 GI:50974358
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 896)
 Shin, J.
 Testis cDNA library
 Unpublished (2004)
 CONTACT: JiHyee Shin
 Division of Animal Genetic Engineering
 School of Agricultural Biotechnology, Seoul National University
 San 56-1, Shillim-dong, Gwanak-gu, Seoul, 151-742, Korea
 Tel: +82-31-479-7557
 Fax: +82-31-479-7550
 Email: paeclloj@hanna1.net
 Seq primer: 73.
 Location/Qualifiers
 1..896
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Korean Native chicken"
 /db_xref="taxon:9031"
 /sex="Male"
 /tissue_type="Testis"
 /dev_stage="26weeks"
 /clone_1lb="Testis cDNA library"
 /note="Vector: Uni-ZAP XR vector; Site 1: XhoI; Site 2:
 EcoRI; The libraries were synthesized using the
 ZAP(R)-cDNA synthesis method (Stratagene) cDNA was
 prepared, size-fractionated and inserted into the Uni-ZAP

FEATURES
 source

XR vector using XhoI linker-primer and EcoRI adaptor.
After in vivo excision with E. coli strain SOLR, cDNA
libraries from testis contain each insert size range of
0.5-3kb (n=20), and 0.5-3.5kb (n=20)."

ORIGIN

Alignment Scores:

| Pred. No.: | 2,71e-110 | Length: | 896 |
|------------------------|-----------|---------------|-----|
| Score: | 1158.00 | Matches: | 240 |
| Percent Similarity: | 81.59% | Conservative: | 17 |
| Best Local Similarity: | 76.19% | Mismatches: | 31 |
| Query Match: | 52.49% | Indels: | 27 |
| DB: | 7 | Gaps: | 2 |

US-10-798-532-4 (1-419) x CO773091 (1-896)

```

QY 46 ProGlySer-Propio-----GlyAlaLeuArgGlnPheLeuProGlyThrTh 62
DB 11 CCTGGNACCTCCACCCGGTGGCGNCCGGCTTAGAACTAGATCCCGGGCTGCAG 70
QY 62 rglYAlaAlaSerAlaLaGluTyrGlyPheArgGlyGluArgAlaAlaLeuGlu 82
DB 71 GAATTCGGCCGAGGC----- 86
QY 82 nleuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSerPr 102
DB 87 -----CGCACATCTGTACTGTGACGGCGTTCACCAACACATCTCGGC 133
QY 102 oAenValProSerArgCyLeuLeuLeuGluLeuLeuAspValSerValSerGluLeu 122
DB 134 CAACGGCCGGTCCCGCTGCTGCTGCTGAGTGTGATATCAGCGTGTGAGTGCCT 193
QY 122 uleuTyrSerSerHisGlnGlyCySerMetTrpMetIleGlnHisCyAlaArgAspVa 142
DB 194 GCTGACCTCCAGCACAGGCGCTGCTCACTGATGATGATCCACACTGCGCCGACAG 253
QY 142 lleuGlnAlaLeuAlaPheLeuHisIleGluGlyTyrValHisAlaAspLeuYsProAr 162
DB 254 GCTGGAAGCCCTGGCTTCTCTGCAACCAAGGCTACGTCCAGCAGACCTGAAGCAG 313
QY 162 gAenIleLeuTrpSerAlaGluAsnGluCyPheLeuLeuIleAspPheGlyLeuSerPh 182
DB 314 CAACATCTGTGAGAGCGCCAGAGAGAGTCTTAAAGCTCATGACTTGGACTTGA 373
QY 182 elySGluGlyAsnGlyAspValYsTyrIleGlnTrpAspGlyTyrArgAlaProGlu 202
DB 374 CAAGAAGGGGAATCAGATGTGAATATATTAACAACAGAGGGGTATCGGGCTC 433
QY 202 agluLeuGlnAsnCyLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCyThrSerAl 222
DB 434 AGAAGTGCAGATCTGCTTACAGACAGGAGGCTCCAGAGTGAAGACGAATGTAC 493
QY 222 aValaPheLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLysLeu 242
DB 494 TGTGATCTGTGAGTCTGGGATGTGTTTAACTGAAATGTTCTCAGAAATGAATG 553
QY 242 shiSThrValArgSerGlnGluTrpYsaIaAsnSerSerAlaIleIleAspHisIlePh 262
DB 554 ACATACAGTCCATCTCAGAAATGAAGACAAACAGTTCGCATATCGATCGCATTTT 613
QY 262 eAlaSer-LysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIle 282
DB 614 TGCATGAAAGGGGGTGTGTTAATTCAGCCATTCAGCTTATCCTCAGAGACTTAT 673
QY 282 ySerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCySerP 302
DB 674 AAAGCATGTCTTATGACCAAGCAAGCAGACCTCTGCTGAAAGGCTTATGACGCC 733
QY 302 roPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrPro 322
DB 734 CATTTTCAGCATTCCTTGTGCTCCCATYGAAGATTTGGATGCTCCCAAGCCCTG 793
QY 322 alLeuArgLeu-LeuAsnValLeuAspAspTyrLeuGlyAsn-GluGluGluTyrG 341

```

```

DB 794 TCCTAGAGGCTGCTGAATGTTCTTAGACGATGCTTCTGCAGAGTGAAGAAATACGA 853
QY 341 uAspValValGlu-AspValYsGluGluCyGln 352
DB 854 AGAATCCCTGAGACATAGAGGAGAGATGTGCAG 888

```

RESULT 9

CN232453

LOCUS CN232453 781 bp mRNA linear EST 09-APR-2004

DEFINITION WLB074D11.ab1 Wtestis Gallus gallus cDNA 5', mRNA sequence.

ACCESSION CN232453

VERSION CN232453.1 GI:46336197

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 781)

REFERENCE Savolainen, P., Fitzsimmons, C.J., Arvestad, L., Andersson, L. and

Lundberg, U.

EST analysis of brain and testis cDNA libraries from White Leghorn

and Red Jungle Fowl

Unpublished (2004)

CONTACT: Peter Savolainen

Department of Biotechnology

Royal Institute of Technology, KTH

SE-106 91 Stockholm, SWEDEN

Tel: +46 (0) 8 5537 8481

Fax: +46 (0) 8 5537 8335

Email: Peter.Savolainen@biotech.kth.se

Seq primer: M13 reverse primer.

Location/Qualifiers

1..781

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn"

/db_xref="taxon:9031"

/sex="male"

/lab_host="ElectroMAX DH10B (Invitrogen)"

/clone_1ib="Wtestis"

/note="Organ: testis; Vector: pSPORT-1; Site 1: Hind III;

Site 2: EcoRI; The cDNA libraries were created with the

Superscript Plasmid System (Invitrogen)."

ORIGIN

Alignment Scores:

| Pred. No.: | 9.14e-108 | Length: | 781 |
|------------------------|---|---------------|-----|
| Score: | 1133.00 <td>Matches:</td> <td>225</td> | Matches: | 225 |
| Percent Similarity: | 93.10% <td>Conservative:</td> <td>18</td> | Conservative: | 18 |
| Best Local Similarity: | 86.21% <td>Mismatches:</td> <td>3</td> | Mismatches: | 3 |
| Query Match: | 51.36% <td>Indels:</td> <td>0</td> | Indels: | 0 |
| DB: | 7 | Gaps: | 0 |

US-10-798-532-4 (1-419) x CN232453 (1-781)

```

QY 109 leuLeuLeuGlnLeuLeuAspValSerValSerGluLeuLeuTyrSerSerHisGln 128
DB 2 CTGCTGCTGAGACTGCTGATATTCAGCGTGTGAGCTGCTGCTCAGTCCGCAACG 61
QY 129 GLYCySerMetTrpMetIleGlnHisCySaIaArgAspValLeuGluAlaLeuAlaPhe 148
DB 62 GGTGTCTCATGTGATGATTCAGACATGCGCCCGACAGTCTTGAAAGCCCTGGCTTC 121
QY 149 leuHisGluGlyTyrValHisAlaAspLeuYsProArgAsnIleLeuTrpSerAla 168
DB 122 CTGCACCAACAAGGGCTAGTGCAGACAGCTGAAGCCACCAATCATCTGTGAGCGCG 181
QY 169 GluAsnGluCyPheLeuLeuLeuIleAspPheGlyLeuSerPheLeuGluGlyAsnGln 188
DB 182 GAGAGAGAGTGTCTTAAAGCTCATGACTTGTGACTTAAAGAGGGGAGATCAGAT 241

```

QY 189 VallyeTyrlleGlnThrAepGlyTyraGalaProGluAlaGluLeuGlnAncyAseu 208
DB 242 GTGAATATATATCAACAAGCGGATCCGGCTCCAGAGGACAGAACTCGAATCCCTTA 301
QY 209 AAG 228
DB 302 GCACAGGAG 361
QY 229 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 248
DB 362 GGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
QY 249 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
DB 422 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 481
QY 269 AAG 288
DB 482 AATTCAG 541
QY 289 ProSerArgArgIleProAlaGluMetAlaLeuCySerProPheSerIleProPhe 308
DB 542 CAAGGCAAG 601
QY 309 AAG 328
DB 602 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661
QY 329 LeuAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 348
DB 662 CTAAAGGAG 719
QY 349 GAG 368
DB 720 GAG 778
QY 369 Arg 369
DB 779 AAA 781
RESULT 10
BUS41306 932 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT_10327419 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6572970
DEFINITION 5', mRNA sequence.
ACCESSION BUS41306
VERSION BUS41306.1 GI:22851747
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 932)
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL Unpublished (1999)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgaops-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LINC2765 row: 0 column: 18
High quality sequence stop: 631.
Location/Qualifiers
1..932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6572970"

/cissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

| Pred. No.: | 6,3e-105 | Length: | 932 |
|------------------------|----------|---------------|-----|
| Score: | 1107.00 | Matches: | 227 |
| Percent Similarity: | 87.22% | Conservative: | 5 |
| Best Local Similarity: | 85.34% | Mismatches: | 18 |
| Query Match: | 50.18% | Indels: | 16 |
| DB: | 5 | Gaps: | 2 |

US-10-798-532-4 (1-419) x BUS41306 (1-932)

QY 1 MetAlaGlySerGlyCyAlaIleProGlyValaGluProProArgPheLeuGluAlaPheGly 20
DB 149 ATGGCGGAGATCCGGCTGGCGCTGGAGCGCGCGCTTTCTGAGCGCTTCGAG 208
QY 21 ArgLeuTrpGlnValaGlnSerArgLeuGlySerGlySerAlaSerValTyraVal 40
DB 209 CGGCTGTGGCAGGTACAGAGCGGTCTGGGTACGGCTCTCCGCTCGGTATGGGTT 268
QY 41 ArgCySerGlyAspProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
DB 269 CGCTCTGGCGCAACCTTGCTCGCGCGCGCGCGCTTTCGAGAGAGAGAGAGAGAG 328
QY 61 ThrThrGlyAlaAlaSerAlaSerAlaGluTyryGlyPheArgGlyGluAlaAlaLeu 80
DB 329 ACCACCGGCGGCGCGCTGCGCGCGCGCGAGATGATTCGCAAGAGAGAGAGAGAG 388
QY 81 GlnGlnLeuGlnGlyAlaArgAsnIleValThrLeuTyryGlyValPheThrIleHisPhe 100
DB 389 GAACAGTTCACAGGTCACAGAAACATCGACCTTTGTATGAGATTTTACATCCACTT 448
QY 101 SerProAsnValProSerArgCyAlaLeuLeuGlnLeuLeuAspValSerValSerGly 120
DB 449 TCTCCAAATGTCCATCAAGCTGTGTGCTTGAATCTCGAGATGATGATGATGATG 508
QY 121 LeuLeuLeuTyrySerSerHisGlnGlyCySerMetTrpMetIleGlnHisCyAlaAlaArg 140
DB 509 TTGCTCTTATATTCAGTCACAGCGCTGTTCATGATGATGATGATGATGATGATG 568
QY 141 AspValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyryValHisAlaAspLeuTy 160
DB 569 GATGTTTGGAGCGCTTCCTTTCTTCATCATGAGGCGTATGTCAGCGGAGCTCCCA 628
QY 160 sProArgAsnIleLeuTrpSerAlaGluAsnGlyCyPheLeuLeuIleAspPheGlyLe 180
DB 629 ACCAGGTAACATATGTGAGATGCAAGAAATGAATGAATGAATGAATGAATGAATG 688
QY 180 u-SerPheLeuGlnGlyLeuGlnAspValTyryIleGlnThrAspGlyTyraGalaP 200
DB 689 TTAGCTTCAAAGAAAGCAATCGAGTGAAGTATATTCAGACAGAGAGAGAGAGAG 748
QY 200 rGlu-AlaGluLeuGlnAsnCySerAlaGlnAla-GlyLeuGlnSerAspTrpGlyCy 219
DB 749 CAGAAACAGAAATTCGAAATATGCTTGCGCCAGCGTGGCTCGCAAGATGATCAAAATG 808
QY 219 sThrSer-AlaValAsp-LeuTrpSer-----LeuGlyIleIleLeuLeuGlnMetPhe 236
DB 809 TACCTTCAGAGTGTGATCTTGAGAGCCCTAAGAAATCATTTTACCTGAAATATGTTCTC 868
QY 237 SerGlyMetLeuLeuLeuHisThrValArgSerGlnGluTrpValAsnSerSerAla 256

DB 869 CAGGAGATG-----GAAACTGGAACCTACCGGTCCAGA 904

QY 257 ||lelle 258

DB 905 ATCTTC 910

RESULT 11

LOCUS BU385761 717 bp mRNA linear EST 28-NOV-2002

DEFINITION 603582650F1 CSBQCHN75 Gallus gallus cDNA clone CHEST533d13 5', mRNA sequence.

ACCESSION BU385761

VERSION BU385761.1 GI:25893762

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

AUTHORS Boardman, P.E., Sanz-Esguero, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

TITLE A Comprehensive Collection of Chicken cDNAs

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE 22335534

PUBMED 12445392

COMMENT Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1. 717

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="CHEST533d13"

/dev_stage="36"

/lab_host="DH10B"

/clone_1lb="CSBQCHN75"

/note="Organ: trunks; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS 1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Alignment Scores:

Pred. No.: 5.41e-105 Length: 717

Score: 1106.00 Matches: 210

Percent Similarity: 93.72% Conservative: 14

Best Local Similarity: 87.87% Mismatches: 14

Query Match: 50.14% Indels: 1

DB: 5 Gaps: 0

US-10-798-532-4 (1-419) x BU385761 (1-717)

QY 134 MetTlGlnHicCyaaIaargAaPvaIlLeuGlaAlaLeuAlaPheUehiShiSglngly 153

DB 3 ATGATCCAGACACTGCGCCGCGACGTCCTGGAAGCCCTTGCCCTTGACACCAAGGC 62

QY 154 TyrValHisaIaAspLeuIyPserArgAenileLeuTPserAlaGluuAsnGluCysPhe 173

DB 63 TACGTGACGACGACCTGAGCAAGCAACATCTCTGTGAGCGCGAGAGAGCTTT 122

QY 174 IyLeuIleIaaspPheGlyLeuSerPheIySglnglyIaangIaaspValIyStryllegln 193

DB 123 AAGCTCATTTAGCTTTGACTTGTAGTTCAAGAGGGGAATTCAGAGATGTAAATATATTCAA 182

QY 194 ThrAaPGLyTyrArgAlaProGluAlaGluuGlnaAnCyLeuAlaGlnAlaGlyLeu 213

DB 183 ACAGAGGGGTATCGGCTTCAGAGGCAAGACTGCAAACTGCTAGCAAGAGGCTC 242

QY 214 GlnSerAaPThrGlnCySThrSerAlaValAspLeuTPserLeuGlylleleleuLeu 233

DB 243 CAGAGTGAACGGAATGTATCTCTCTGTGATCTGTGAGTCTGGAAATGTTTACTCG 302

QY 234 GluMetPheSerGlyMetIySglnglyIyStryValArgSerGlnGluTrrlyAlaAaen 253

DB 303 GAAATGTTCTCAGAAATGAACCTGAACATACAGTCAATCTCAGGAATGMAAGCAAAAC 362

QY 254 SerSerAlaIleIleIaaspHisIlePheIaSerIyAlaValaAlaAlaIlePro 273

DB 363 AATTCTGCATCATCATGATCGCAATTTTGCAGTGAAGGGGTGTAAATTCAGCCATTCCA 422

QY 274 AlaTyrHisIleuArgAspLeuIleIySserMetLeuHisAaspPProSerArgArgIle 293

DB 423 GCTTATACCTACGACGACCTTATTAAGCATGCTTCAATGTGACCAAGCAAGAGGC 482

QY 294 ProIaGluMetAlaLeuCySerProPhePheSerIleProPheAlaProHisIleGlu 313

DB 483 TCTGCTGAAGAAAGCTTTATGAGCCCATCTTCACATATCCCTTGTGCTCCCATATGAA 542

QY 314 AspLeuValMetLeuProThrProValIleuArgLeuLeuAsnValLeuAspAspStry 333

DB 543 GATTGTGGATGCTCCCAAGCGCTGTGCTGAGCGTGAATGTTCTAAGCATCTTCT 602

QY 334 LeuGlyIyAsnGluGluTyrGluAspValIyGluAspValIyGluCysGlnIy 353

DB 603 CTGCAGAGTGAAGAAAGATACGAATATCTCGGAAGACATAGAGAGAGTCTCAGAAA 662

QY 354 TyrGlyProValIySerLeuLeuValProIySgluAanProGlyIyArglyGln 371

DB 663 TATGAGACCGGTGTTCTTCTTGTATTCGAAAGAAACATCTGTAAAGGCCAA 717

RESULT 12

LOCUS BE895119 899 bp mRNA linear EST 20-OCT-2000

DEFINITION 601436041F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920939 5', mRNA sequence.

ACCESSION BE895119

VERSION BE895119.1 GI:10358193

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC/DC/DTP

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILML at: <http://image.llnl.gov>

http://image.llnl.gov

Plate: ILMA9753 row: b column: 12

High quality sequence stop: 626.

Location/Qualifiers

SOURCE

1. .899
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:392093"
 /issue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies.."

ORIGIN

Alignment Scores:

| Pred. No.: | 2,83e-103 | Length: | 899 |
|------------------------|-----------|---------------|-----|
| Score: | 1091.00 | Matches: | 210 |
| Percent Similarity: | 99.53% | Conservative: | 0 |
| Best Local Similarity: | 99.53% | Mismatches: | 0 |
| Query Match: | 49.46% | Indels: | 1 |
| DB: | 2 | Gaps: | 0 |

US-10-798-532-4 (1-419) x BE895119 (1-899)

QY 113 LeuLeuAspValSerValSerGluLeuLeuTyrSerSerHisGlnGlyCysSerMet 132
 |||
 DB 1 CTCCTGGAGTGTGAGTGTTCGGAATTCCTTATATTCAGTCACAGGGTGTTCAG 60
 QY 133 TTPMeTIIeGlnHisCysAlaArgAspValLeuGlnAlaLeuAlaPheLeuHisGlu 152
 |||
 DB 61 TGGATATACAGCATTTGTGCCGAGATGTTTGGAGGCCCTTGTCTTCATCATGAG 120
 QY 153 GATYTYValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGlyCys 172
 |||
 DB 121 GGGTATGTCCATCGGAGCTCAACCAAGCAATATTTGGAGTGCAGAGATGATGT 180
 QY 173 PheLysLeuIleAspPheGlyLeuSerPheLysGlnGlyAsnGlnAspValLysTyrIle 192
 |||
 DB 181 TTTAACTATTCATCTTGGACTTACCTTCAGAAAGGCAATCAGATGTAAGTATAT 240
 QY 193 GlnThrAspGlyTyrTrpAlaProGlnAlaGlnLeuGlnAsnGlyLeuAlaGly 212
 |||
 DB 241 CAGACGACGGGTATGGGCTCCAGAGCAAGATTCGAAATTCCTTGGCCCGCTGGC 300
 QY 213 LeuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeu 232
 |||
 DB 301 CTGCAGAGTATCAAAATGTACTCAGCTGTGATCTGGAGCTTAGGAATCATTTTA 360
 QY 233 LeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTyrLysAla 252
 |||
 DB 361 CTGGAAATGTTCTCAGGAATGAATGAACATACATACAGTCTCAGGAATGGAAGGCA 420
 QY 253 AsnSerSerAlaIleIleAspHisIlePheAlaSerTyrValAlaAsnAlaIle 272
 |||
 DB 421 AACAGTTCGCTATATGATCATATTTGCCAGTAAGCAGTGGGAATGCGCAAT 480
 QY 273 ProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSerArgArg 292
 |||
 DB 481 CCGAGCTATCACTTAGAGAGCTTATCAAAAGCATGTTTATGATATCAACAGAGAG 540
 QY 293 IleProAlaGluMetAlaLeuCysSerProPhePheSerIleProPheAlaProHisIle 312
 |||
 DB 541 ATTCCTGCTGAATGGCATTTGTCCAGGCCCATCTTTTAGCATTCCTTTTGGCCCTCATATT 600
 QY 313 GluAspLeuValMetLeu-ProThrProVal 322
 |||
 DB 601 GAAGATCTGATCATGCTTCCACTCAGTT 631

RESULT 13

BI562420 772 bp mRNA linear EST 05-SEP-2001
 LOCUS BI562420
 DEFINITION 60326316P1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5298674 5',
 mRNA sequence.

ACCESSION

BI562420
 VERSION BI562420.1 GI:15449747
 KEYWORDS EST.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

NIH-MGC http://mhc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1993)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LLM11755 row: P column: 03
 High quality sequence stop: 727.

FEATURES

source

1. .772
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5298674"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_97"
 /note="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
 (gtcgaq); Oligo-dT primed using primer
 5'-TTTTTTTTTTTTTTVN-3', size-selected for average
 insert size 2.2 kb and normalized to ROT 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: this is
 a NIH_MGC Library."

ORIGIN

Alignment Scores:

| Pred. No.: | 4,06e-102 | Length: | 772 |
|------------------------|-----------|---------------|-----|
| Score: | 1079.00 | Matches: | 212 |
| Percent Similarity: | 96.82% | Conservative: | 1 |
| Best Local Similarity: | 96.35% | Mismatches: | 5 |
| Query Match: | 48.91% | Indels: | 3 |
| DB: | 4 | Gaps: | 1 |

US-10-798-532-4 (1-419) x BI562420 (1-772)

QY 92 LeuTyrGlyValPheThrIleHisPheSerProAsnValProSerArgCysLeuLeu 111
 |||
 DB 108 TTTGATGGAGTGTTCACATCACTTTCTCCAAATGTCATCACGCTGTCTGTGCTT 167
 QY 112 GluLeuLeuAspValSerValSerGluLeuLeuTyrSerSerHisGlnGlyCysSer 131
 |||
 DB 168 GAATCTCTGATGTACGTGTTTCGGAAATGCTTATATTCAGTACACCAAGGTGTTCC 227
 QY 132 MetTTPMeTIIeGlnHisCysAlaArgAspValLeuGlnAlaLeuAlaPheLeuHisGlu 151
 |||
 DB 228 ATCTGATGATACAGCATTTGTCCGAGATGTTTGGAGGCCCTTGTCTTTTCAATCAT 287
 QY 152 GluGlyTyrValHisAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGlu 171
 |||
 DB 288 GAGGCTATGTCCATGCGGACCTCAACCAAGTACATATTTGTGAGTGCAGAGATGAA 347
 QY 172 CysPheLysLeuIleAspPheGlyLeuSerPheLysGlnGlyAsnGlnAspValLysTyr 191
 |||
 DB 348 TGTTTAACTATTCATCTTGGACTTACCTTCAGAAAGGCAATCAGATGTAAAGTAT 407

192 IleglnthAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAla 211
408 ATTCAAGACAGACGGGATCGGGCTCCAGAGACGAAATTCGAAATTCCTGGCCAGGCT 467
212 GtyleuGlnSerAspThrGluCysThrSerAlaValAspLeuTrpSerLeuGlyIleIle 231
468 GGCCTGCAGAGATGATACAAATGTACTCAGCTGTGATCTGTGGAGCCTAGGAATCATTT 527
232 LeuLeuGlnMetPheSerGlyMetLeuLeuYHisThrValArgSerGlnGluTrpLys 251
528 TTACTGGAAATGTTCTCAGGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 587
252 AlaAsnSerSerAlaIleIleAspHisIle-PheAlaSerLysAlaValAlaAsnAla 271
588 GCAAAACAGTTCTGCTATTATGATCATATTATTTGGCAATGAAGCAGTGATGATCCGC 647
271 aAlleProAlaTyrHisLeuArgAspLeuIleLysSerMetLeuHisAspAspProSer 291
648 AATTCAGAGCTTACACTTAAGAGACCTTAACAAAGACATGCTTCATGATGATCAAGCAG 707
291 gAtgIleProAlaGluMetAlaLeuGlnCysSerProPhePheSerIleProPheAlaPro 310
708 AAGAATTCCTGCTGAATGACATGTGCAG-CCATTCTTAGCA--TCCCTTGGCCCT 761

RESULT 14
BU205357 794 bp mRNA linear EST 25-NOV-2002
LOCUS 603104569F1 CSEQCHN03 Gallus gallus cDNA clone CHEST41118 5', mRNA
DEFINITION sequence.
ACCESSION BU205357
VERSION BU205357.1 GI:25370848
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 794)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Boesch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken CDNAS
Curr. Biol. 12 (22), 1965-1969 (2002)
2235534
12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. 794
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, HiseX"
/db_xref="taxon:9031"
/clone="CHEST41118"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_idb="CSEQCHN03"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores:

| Pred. No.: | 1,11e-101 | Length: | 794 |
|------------------------|-----------|---------------|-----|
| Score: | 1075.00 | Matches: | 221 |
| Percent Similarity: | 89.10% | Conservative: | 16 |
| Best Local Similarity: | 83.08% | Mismatches: | 27 |
| Query Match: | 48.73% | Indels: | 4 |
| DB: | 5 | Gaps: | 0 |

US-10-798-532-4 (1-419) x BU205357 (1-794)

82 GtlnleuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPheSer 101
1 CAGCTCCGGGGGACCGCAACATCGTACTGTACGGCGGTGTCCACCAACCACTACTCG 60
102 ProAlaValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGluLeu 121
61 GCCAACGGCCCGCTCCGCTGCTGCTGAGCTGTGGATATACGGGTCTTAGCTG 120
122 LeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArgAsp 141
121 CTGCTGCACTCCAGCAACAGAGTCTGCTCCATGATGATGATCAGACACTGCCGCCGAC 180
142 ValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLysPro 161
181 GTCTGGAAGCCCTGGCTGCTCTGACACCAAAAGGCTACGTACAGCAGCACTGAAGCCA 240
162 ArgAsnIleLeuTrpSerArgAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeuSer 181
241 CCGCAACATCTGTGAGAGCGCGGAGAGAGAGTCTTAAGCTCATGACTTTGGACTTTGAC 300
182 Phe-LysGlnGlyAsnGlnAspValLysTyrIleGlnthAspGlyTyrArgAlaProGln 201
301 TTCAAAAGGGGGAATCAGAGATGTAATATATTTCAACACAGCGGATTCGGGCTCCAGA 360
201 uAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThrSe 221
361 GGCAGAGCTGCAAGACTGCTGCTAGCAGAGGAGGCTCCAGAGTACAGCAAGATGACTTC 420
221 rAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLysIle 241
421 TCGCTGGATCTGTGAGTCTGGAGATTTGTTTACTGGAAATGTTCTCAGGAATGAAACT 480
241 uLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIle 261
481 GAAACATACAGTCCATATCCAGGAATGAAAGCAACAGTTTGCATCATCATGATCGCAT 540
261 ePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeuIle 281
541 TTTTCCAGTGAAGGGGTGTATATCCAGCCATTCAGATTATCACTCAGAGACTTAT 600
281 eLysSerMetLeuHisAspAsp-ProSerArgArgIleProAlaGluMetAlaLeuCys 301
601 TAAAGACATGCTTCAATTGGAACCAAGCAAAAGAGCTCTGTGAAAGGCTTATATGCA 660
301 eProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThrP 321
661 GCCCATCTTCAAGCATTCCTTGCTGCC-CAATATGAAGATTGTGATGCTCCCAAG-C 718
321 rValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlnGlyAsnGlnGlnGlyTyr 341
719 CTGTCTGAGGCTGCTGATATGTTCTAAGCGAAGGTCTCTGCAAGTGAAGAAATAACG 778
341 LuAspValValGlu 345
779 AGAATATCTGGGAA 792

RESULT 15

CF593363 834 bp mRNA linear EST 26-SBP-2003
 LOCUS AGENCOURT 15622279 NIH MGC 147 Homo sapiens cDNA clone
 DEFINITION IMAGE:10531268 5', mRNA sequence.
 ACCESSION CF593363
 VERSION CF593363.1 GI:36346826
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 834)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contract: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDM621 row: m column: 05
 High quality sequence scop: 660.
 Location/Qualifiers

FEATURES

1..834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30531268"
 /issue_type="Human Placenta"
 /lab_host="DH10B Tona"
 /clone_lib="NIH MGC 147"
 /note="Organ: Placenta; Vector: pBluescript; Site 1:
 all-XhoI; Site 2: BamI; oligo-dT primed using primer
 5'-TTTTTTTTTTTTTTVN-3', size-selected for average
 insert size 2.3 kb and normalized to R07 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: This is
 a NIH-MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 3,54e-101 Length: 834
 Score: 1070.50 Matches: 213
 Percent Similarity: 95.13% Conservative: 2
 Best Local Similarity: 94.25% Mismatches: 5
 Query Match: 48.53% Indels: 6
 DB: 7 Gaps: 1

US-10-798-532-4 (1-419) x CF593363 (1-834)

QY 1 MetAlaGlySerGlyCySaLaTrpGlyAlaGluProArgPheLeuGluAlaPheGly 20
 DB 155 ATGGCGGATCCGGCTGGCGCTGGGGCGGAGCCCGCGTTTCTGGAGGCTTCGGG 214
 QY 21 ArgLeuTrpGlnValInserArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB 215 CGGCTGTGGAGTACAGAGCCGCTGGGTAGCGGCTCCCTCGGCTGATCGGGTT 274
 QY 41 ArgCyAGlyAanProGlySerProGlyAlaLeuArgGlnPheLeuProGly 60
 DB 275 CGCTGCTGGGCAACCTGCTGCGCCCGGCGCCCTCAAGCACTTCTTGGCCGCAAGA 334
 QY 61 ThrThGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgIysGluArgAlaAlaLeu 80

DB 335 ACCACCGGGGCTGCGGCTCTGCGCCGAGATGCTTCCGAAAGAGAGGGCGGCTG 394
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB 395 GAAACAGTTCAGAGGTCACAGAAACATCGTACTTGTATGAGAGTGTATACATCCACTT 454
 QY 101 SerProAsnValProSerArgCysLeuLeuGlnLeuLeuAspValSerValSerGlu 120
 DB 455 TCTCCAAATGTGTCATCAGCTGCTGTGTGCTTGAATCTCGAGATGATGATGATGATG 514
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 515 TTGCTCTTATATTCAGTCACAGGAGTGTTCATGATGATGATGATGATGATGATGATG 574
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisIleGlnGlyTyrValHisAlaAspLeuLys 160
 DB 575 GATGTTTGGAGGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 634
 QY 161 ProArgAsnIleLeuTyrSerArgIleGlnGlnGlyCysPheLysLeuIleAspPheGly-Le 180
 DB 635 CCACGTAACATATGTGTGAGATGACAGAAATGATTTTAACTCATTCATTCATTCGACTT 694
 QY 180 userPheLysGluGlyAanGlnAspValLysTyrIleGlnThrAspGlyTyrArg-Alap 200
 DB 695 AGCTTCAAAAGAAACAAATCAGAGATGTAAAGTATTCAGACAGAGGGATATCGGGCTC 754
 QY 200 roGluAlaGluLeuGln-AsnCyseuAla---GlnAlaGly-LeuGlnSerAspThrG1 218
 DB 755 CAGAAAGCAAAATTCGAAAAATTCCTTGGGCCCCAGGCTTGGCCCTGCAAGATGATACGA 814
 QY 218 u-CysThr 220
 DB 815 AATGTACC 822

Search completed: November 29, 2004, 13:27:52
 Job time : 3168.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 502 Seconds
(without alignments)
4381.489 Million cell updates/sec

Title: US-10-798-532-4
Perfect score: 2206
Sequence: 1 MAGSGCAGABPPRFLEAFG.....VATFYPPLSAVKRGYLYQTLL 419

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 413486 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=rlp
-Q/cg12.1/USPTO.epool_p/US10798532/runat_29112004.112452.15882/app_query.fasta.1.1166
-DB=N Geneseq.23Sep04 -QPM=fastap -SUFPX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNIT8=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALLIN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=ptc -NORM=exp -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10798532.0CGN_1.1.886.0runat_29112004.112452.15882 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: N_Geneseq.23Sep04.*
2: geneseqn1980s.*
3: geneseqn1990s.*
4: geneseqn2000s.*
5: geneseqn2001as.*
6: geneseqn2001bs.*
7: geneseqn2002as.*
8: geneseqn2002bs.*
9: geneseqn2003as.*
10: geneseqn2003bs.*
11: geneseqn2003cs.*
12: geneseqn2003ds.*
12: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | |
|------------|-------|--------------------------|--------------------|
| Result No. | Score | Query Match Length DB ID | Description |
| 1 | 2206 | 100.0 1260 3 | AAZ51356 Human tra |
| 2 | 2203 | 99.9 1260 4 | AAZ51355 Human wil |
| 3 | 2195 | 99.5 1260 4 | AAI6829 Human pro |
| 4 | 2195 | 99.5 2008 6 | ABQ9383 Human cod |
| 5 | 2195 | 99.5 2622 4 | AAI6828 Human pro |
| 6 | 2195 | 99.5 4065 4 | AAH72714 Human cer |

| | | | | | | |
|----|--------|------|------|----|-----------|---------------------|
| 7 | 2195 | 99.5 | 4065 | 5 | ADL45238 | Adl45238 Human ova |
| 8 | 2188 | 99.2 | 1260 | 4 | AA506734 | AA506734 Polynucle |
| 9 | 2188 | 99.2 | 1824 | 4 | AAI18818 | AAI18818 Human kin |
| 10 | 2183 | 99.0 | 3244 | 10 | AB52857 | AB52857 Primary r |
| 11 | 2042.5 | 92.6 | 3598 | 12 | ADQ24401 | ADQ24401 Human bof |
| 12 | 789 | 95.8 | 529 | 4 | AAH70764 | AAH70764 Human cer |
| 13 | 766.5 | 94.7 | 722 | 3 | AAA02535 | AAA02535 Human col |
| 14 | 729 | 93.0 | 470 | 4 | AAH71966 | AAH71966 Human cer |
| 15 | 636 | 28.8 | 461 | 4 | AAH69101 | AAH69101 Human cer |
| 16 | 559.5 | 25.4 | 411 | 6 | ABQ99225 | ABQ99225 Human ORF |
| 17 | 361 | 16.4 | 231 | 3 | AAH01259 | AAH01259 Human col |
| 18 | 311.5 | 14.1 | 419 | 5 | AD173740 | AD173740 Human ova |
| 19 | 311.5 | 14.1 | 419 | 5 | AD167345 | AD167345 Human ova |
| 20 | 284 | 12.9 | 546 | 6 | ABQ53127 | ABQ53127 Oligonuc1 |
| 21 | 284 | 12.9 | 546 | 6 | ABQ53126 | ABQ53126 Oligonuc1 |
| 22 | 284 | 12.9 | 547 | 6 | ABQ44465 | ABQ44465 Oligonuc1 |
| 23 | 284 | 12.9 | 547 | 6 | ABQ44464 | ABQ44464 Oligonuc1 |
| 24 | 255 | 11.6 | 546 | 6 | ABQ53129 | ABQ53129 Oligonuc1 |
| 25 | 255 | 11.6 | 546 | 6 | ABQ53128 | ABQ53128 Oligonuc1 |
| 26 | 255 | 11.6 | 547 | 6 | ABQ44463 | ABQ44463 Oligonuc1 |
| 27 | 255 | 11.6 | 547 | 6 | ABQ44462 | ABQ44462 Oligonuc1 |
| 28 | 250 | 11.3 | 442 | 9 | ACH28557 | ACH28557 Human adu |
| 29 | 249 | 11.3 | 1065 | 10 | ADD21399 | ADD21399 Golden ha |
| 30 | 246.5 | 11.2 | 1400 | 10 | ADD21407 | ADD21407 Rat cycl1 |
| 31 | 245 | 11.1 | 1708 | 10 | ADD21395 | ADD21395 Mouse cyc |
| 32 | 240.5 | 10.9 | 2472 | 3 | ACA45716 | ACA45716 Arabidops |
| 33 | 240 | 10.9 | 1297 | 10 | ADD21393 | ADD21393 Human cyc |
| 34 | 240 | 10.9 | 2213 | 6 | ABK84073 | ABK84073 Human cDN |
| 35 | 240 | 10.9 | 2213 | 10 | ADD21403 | ADD21403 Chinese h |
| 36 | 240 | 10.9 | 2213 | 11 | AD131711 | AD131711 Human cDN |
| 37 | 240 | 10.9 | 2213 | 12 | ADH17138 | ADH17138 Human cdc |
| 38 | 240 | 10.9 | 2328 | 10 | ACC72837 | ACC72837 Human can |
| 39 | 240 | 10.9 | 2734 | 12 | ADQ25034 | ADQ25034 Human sof |
| 40 | 239 | 10.8 | 1650 | 12 | ADD145326 | ADD145326 Rice isop |
| 41 | 237 | 10.7 | 1209 | 10 | ADD21401 | ADD21401 Golden ha |
| 42 | 236.5 | 10.7 | 1240 | 6 | AB188413 | AB188413 Pain regu |
| 43 | 236.5 | 10.7 | 1335 | 2 | AAV71073 | AAV71073 CDK2-gree |
| 44 | 236.5 | 10.7 | 1655 | 3 | AAA09689 | AAA09689 Rat stres |
| 45 | 236 | 10.7 | 1297 | 6 | ABQ60787 | ABQ60787 Human 10C |

ALIGNMENTS

| | | |
|----------|--|---|
| RESULT 1 | AAZ51356 | standard; DNA; 1260 BP. |
| ID | AAZ51356 | |
| XX | AAZ51356; | |
| AC | | |
| XX | | |
| DT | 06-JUN-2000 | (first entry) |
| XX | | |
| DE | Human transdominant mutant serine/threonine kinase KIS (hKIS) gene. | |
| XX | | |
| KW | KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase; | |
| KW | transdominant gene; mutant; cyclin-dependent kinase inhibitor; CKI; p27; | |
| KW | modulator; treatment; cell proliferative disease; vascular disorder; | |
| KW | gene therapy; atherosclerosis; restenosis; ds. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| FH | Key | Location/Qualifiers |
| FT | CDS | 1..1260 |
| FT | | /tag= a |
| FT | | /product= "Human transdominant mutant KIS protein" |
| FT | | /transl_except= (pos:556..561, Asn) |
| FT | misc_feature | 161 |
| FT | | /tag= b |
| FT | | /note= "A to G substitution resulting in a Lys to Arg |
| FT | | change in the amino acid sequence" |
| XX | | |
| PN | WO200011165-A1. | |
| XX | | |
| PD | 02-MAR-2000. | |

XX 20-AUG-1999; 99WO-US018903.
XX
XX 21-AUG-1998; 98US-0097710P.
XX
XX (NABEL/) NABEL G J.
XX (NABEL/) NABEL E G.
XX
XX Nabel GJ, Nabel EG;
XX
XX MPI; 2000-237648/20.
XX P-PSDB; AAY70306.
XX
XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
XX PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
XX PT proliferation.
XX
XX Claim 7, Page 59; 70pp; English.
XX
XX The present DNA sequence encodes a transdominant mutant human KIS (hKIS),
XX CC constructed by site directed mutagenesis. A single nucleotide
XX CC substitution (A to G) results in a lysine to arginine change in the
XX CC protein sequence. hKIS is a serine/threonine kinase that acts as an
XX CC inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
XX CC controls cell proliferation and is localised predominantly in the
XX CC nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
XX CC G1 phase. The hKIS sequences are used to modulate cell proliferation and
XX CC treat cell proliferative and vascular diseases. The polynucleotide
XX CC sequence may be used in gene therapy to treat vascular disorders such as
XX CC restenosis or atherosclerosis
XX
XX Sequence 1260 BP; 310 A; 280 C; 336 G; 334 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1 44e-207 Length: 1260
Score: 2206.00 Matches: 419
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-798-532-4 (1-419) x AA251356 (1-1260)

QY 1 MetAlAGlYserGIcYsaLaTRpGIyAlaGIuPRoPArPhelEuGIuAlaPhelGIy 20
Db 1 ATGGCGGGAATCCGGCTGGCGCTGGGGCGGAGCCGGCGCTTTTCGAGGCGCTTCGGG 60
QY 21 ArgLeuTRpGIuAlaGIuSerArgLeuGIySerGIySerSerAlaSerValTYrArgVal 40
Db 61 CGGCTGTGGAGGTACAGAGCCGCTGGGTAGCGGCTCCCGCTCGGTATCGGGTT 120
QY 41 ArgCYseGIyAnSPrOGlYserPRoGIyAlaLeuArgInPhelEuPRoGIy 60
Db 121 CGCTGTGGCGCAACCTGGCTGGCCCCCGGCGCCTCAGGAGTCTTTCGCCGCAAGA 180
QY 61 ThrThGIyAlaAlaAlaSerAlaAlaGIuTYrGIyPhelArgIySGLuArgAlaLeu 80
Db 181 ACCACCGGGGCTCGGCGCTCTGCGCGAGTATGTTTCGCAAGAAGAGGGCGGCTG 240
QY 81 GluGlnLeuGlnGlnIyHsArgAnIleValThrLeuTYrGIyAlaPhelTrIleHsPhe 100
Db 241 GAACAGTTGGAGGTACAGAAACATCGTACTTTGTATGAGAGTGTTCACATCCACTT 300
QY 101 SerPRoAnValPRoSerArgCYseLeuLeuGIuLeuEuAnSPyAlaSerValSerGIu 120
Db 301 TTCACAAATGTGCATCACGCTGTCTGCTTGAACTCCTGATGTCACTGTTTGGAA 360
QY 121 LeuLeuLeuTYrSerSerHsGIuGIyCYseSerMetTRpMetIleGlnHsCYsaLaArg 140
Db 361 TTGCTCTTATATTCAGTCAACAGGGGTGTTCATGTGATGATACAGATTGCGCCGA 420
QY 141 AspValLeuGIuAlaLeuAlaPheLeuHsHsGIuGIyTYrValHsAlaAspLeuTYr 160
|||

Db 421 GATGTTTTGGAGGCCCTTCCTTTCTTCATCATGAGGCGTATGTCATCGGACCTCAA 480
QY 161 ProArgAnIleLeuTRpSerAlaGIuAnGIuCYsePheYLeuIleAspPheGIyLeu 180
Db 481 CCACGTACATATTTGTGAGTGCAGAGAAATGATGTTTAAATCATTCATTCAGTTCGACTT 540
QY 181 SerPheYLeuGIyAnSPyAlaPheValTYrIleGlnThrAspGIyTYrArgAlaPro 200
Db 541 ACCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGAGGGGATCGGCTCCA 600
QY 201 GluAlaGIuLeuGIuAnSPyAlaGIuAlaGIuIleProAlaTYrHsIleuArgAspLeu 220
Db 601 GAAGCAGAAATTCANAAATGTCTTGCGCCAGCTGGCTGCACAGTGTATACAAATGTAC 660
QY 221 SerAlaValaSPLeuTRpSerLeuGIyIleIleLeuLeuGIuMetPheSerGIyMetIys 240
Db 661 TCAGCTGTTGATCTGTGGAGCCTAGAAATCATTTTACTGGAAATGTTTCAGGAATGAA 720
QY 241 LeuLYHsIleThrValArgSerGIuTYrPlyAlaAnSPSerAlaIleIleAspHs 260
Db 721 CTGAAACATACAGTCAAGATCTCAGAAATGAAAGCAACAGTTCGTATTTGATCAC 780
QY 261 IlePheAlaSerIleAlaValaIleAlaIleProAlaTYrHsIleuArgAspLeu 280
Db 781 ATATTTGCCAGTAAAGCAAGTGTGAATGCCCAATTCAGCTTATCACCTTAGAGACTT 840
QY 281 IleYSerMetLeuHsAspAPrOserArgArgIleProAlaGIuMetAlaLeuCYs 300
Db 841 ATCAAAAGCATCTTCATGATGATCCAGCAAGAAATTCCTCGTGAATGSCATTTGTC 900
QY 301 SerPRoPhePheSerIlePRoPheAlaPRoHsIleGIuAspLeuValMetLeuPRoThr 320
Db 901 AGCCCATTCCTTACATTCCTTTTCCCTCATTTGAAGATCTGTCATGCTTCCACT 960
QY 321 ProValIleuArgLeuLeuAnSPyAlaIleuAspAspTYrLeuGIyAnSPGIuGIuTYr 340
Db 961 CCAAGTCTAAGCTGCTGATGTGCTGATGTGATGATATCTTGGAAATGAAAGAAAT 1020
QY 341 GluAspValaIleGIuAspValaIySGIuGIyCYseGlnTYrGIyProValaIleSerLeu 360
Db 1021 GAAGATGTTGTGAAGATGTAAAGAGGAGTCAAAAATATGACCAAGTGTATCTCTA 1080
QY 361 LeuValPRoLYeGIuAnSPrOGlYArgGIyGlnValPheValGIuTYrAlaAnSPAGIy 380
Db 1081 CTTGTTCCAAAGAAATCTCTGCAAGAGCAAGCTTGTGTGATGATGCAATCTGCT 1140
QY 381 AspSerIyAlaAlaGIuIyLeuLeuTYrArgMetPheAspGIyLYAspPheValaI 400
Db 1141 GATTCAAAGCTGCGCAGAAATTAAGTGAAGAGATGTTGATGGAAAGTGTGTG 1200
QY 401 AlaThrPheTYrPRoLeuSerAlaTYrIyArgGIyTYrLeuTYrGlnThrLeuLeu 419
Db 1201 GCTACATTTACCCCTGAGTGCCTTACAAAGAGGATATCTGTATCAAACTTGCTT 1257
RESULT 2
AAZ51355
ID AAZ51355 standard; DNA; 1260 BP.
XX
AC AAZ51355;
XX
DT 06-JUN-2000 (first entry)
XX
XX Human wild type serine/threonine kinase KIS (hKIS) gene.
DE
XX KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
XX dominant gene; cyclin-dependent kinase inhibitor; CKI; p27; modulator;
XX treatment; cell proliferative disease; vascular disorder; gene therapy;
XX atherosclerosis; restenosis; de.
OS
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1260

```

PT      /tag= a
PT      /product= "Human wild type KIS protein"
PT      /transl_except= (pos:556..561, Asn)
XX
XX      MO200011165-A1.
XX
XX      02-MAR-2000.
XX
XX      20-AUG-1999; 99MO-US018903.
XX
XX      21-AUG-1998; 98US-0097710P.
XX
XX      (NABE/) NABEL G J.
XX      (NABE/) NABEL E G.
XX
XX      Nabel GJ, Nabel EG;
XX
XX      WPI; 2000-237648/20.
XX      P-PSDB; AAY70305.
XX
XX      Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
XX      for inhibiting the cyclin kinase inhibitor p27, and so alter cell
XX      proliferation.
XX
XX      Claim 1; Page 56; 70pp; English.
XX
XX      The present DNA sequence encodes a wild type human KIS (hKIS), isolated
XX      from a human B-cell library, using a yeast two hybrid screening system.
XX      hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
XX      cyclin-dependent kinase inhibitor (CKI), p27. hKIS is a dominant gene
XX      that controls cell proliferation and is localised predominantly in the
XX      nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
XX      G1 phase. It has 99% homology to rat serine/threonine protein kinase KIS.
XX      The hKIS sequences are used to modulate cell proliferation and treat cell
XX      proliferative and vascular diseases. The polynucleotide sequence may be
XX      used in gene therapy to treat vascular disorders such as restenosis or
XX      atherosclerosis
XX
XX      Sequence 1260 BP; 311 A; 280 C; 335 G; 334 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 2,84e-207 Length: 1260
XX      Score: 2203.00 Matches: 418
XX      Percent Similarity: 100.00% Conservative: 1
XX      Best Local Similarity: 99.76% Mismatches: 0
XX      Query Match: 99.86% Indels: 0
XX      DB: 3 Gaps: 0
XX
XX      US-10-798-532-4 (1-419) x AA251355 (1-1260)
XX
QY      1 MetAlaGlySerGlyCysAlaATrPGlYAlaGluProProArhPheLeuGluAlaPheGly 20
DB      1 ATGGCCGGATCCCGGCTGCGCTGGGGCGGAGCCCGCGGTTTCGAGGGCTTCGGG 60
QY      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArGVal 40
DB      61 CGGCTGTGGCAGGTACAGACCCCTCTGGGTAGCGGCTCTCCGCCCTCGGTATCCGGTT 120
QY      41 ArgCysCysGlyValAndProGlySerProProGlyValAlaLeuArgGlnPheLeuProProGly 60
DB      121 CGCTGCTGGCGCAACCTGGCTGCGCCCGCGGCGCTCAAGCAGTTCTTGGCCGACAGA 180
QY      61 ThrThrglyAlaAlaAlaSerAlaAlaGluTYrGlyPheArgGlySerGluArgAlaAlaLeu 80
DB      181 ACCACCGGGGCGGCGCTCTGCGCGGAGTATGTTTCGCAAAAGAGGGCGGCTG 240
QY      81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTYrGlyValPheThrIleHisPhe 100
DB      241 GAACAGTTGAGGGGTCAAGAAACATCGTACCTTTGATGAGATGTTTACATCCACTTT 300
QY      101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
DB      301 TTCCTAAATGTGCATACAGCTGTCTGTGCTTGAACCTCGATGATCAGTGTTCGAA 360

```

```

QY      121 LeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
DB      361 TTGCTCTTATATTCAGTACACAGGGTTGTTTCATGTGATGATACAGATTCGGCCCGA 420
QY      141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTYrValHisAlaAspLeuLys 160
DB      421 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGCTATGTCCATGCGGACCTCAAA 480
QY      161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
DB      481 CCAAGTACATATTTGGAGTGCAGAGATGATGTTTAACTCAATGACTTGGACTT 540
QY      181 SerPheLysGluGlyValAsnGlnAspValLysTYrIleGlnThrAspGlyTYrArgAlaPro 200
DB      541 AGCTTCAAAGAAAGGCAATCAGATGTAAAGTATATTCAGACAGAGCGGATCGGCTCCA 600
QY      201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
DB      601 GAAGCAGAAATTCAAAATTGCTTGGCCCGCTGGCTGCAGAGTATACAGAAATGTACC 660
QY      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
DB      661 TCAGCTGTGATCTGTGAGCTGAGCTAGGAATCATTTTACGGAATGTCTCAGAAATGAAA 720
QY      241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
DB      721 CTGAAACATACAGTACGATCTCAGAAATGAAAGCAACAGTTCGCTTATTTATCAC 780
QY      261 IlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTYrHisIleuArgAspLeu 280
DB      781 ATATTGTCAGTAAGACAGATGTGATGCGCAATTCAGCTTACACTTAAGAGACCTT 840
QY      281 IleLysSerMetLeuHisAspAspProSerArgGlyIleProAlaGluMetAlaLeuCys 300
DB      841 ATCAAAAGATGCTTCATGATATATCCAGCAGAAATTCCTGTCGAATGCAATGTGTC 900
QY      301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB      901 AGCCCATTTTATGATTCCTTTTGGCCCTCATATTTGAAGATCTGTCATGCTCCACT 960
QY      321 ProValLeuArgLeuLeuAsnValLeuAspAspTYrLeuGlyValAsnGluGluTYr 340
DB      961 CCAGTGTCTAAGCTGCTGATGTGCTGATGATGATTTCTTGGAAATGAAGAGAAAT 1020
QY      341 GluAspValValGluAspValLysGlnGlyCysGlnLysTYrGlyProValValSerLeu 360
DB      1021 GAAGATGTTGTGAAGATGTAAAGAGAGTGTCAAAATATGACACAGTGTATCTCTA 1080
QY      361 LeuValProLysGluAsnProGlyValArgGlyGlnValPheValGluTYrAlaAsnAlaGly 380
DB      1081 CTGTGTCAAAGAAATCTCTGCAAGAGCAAGCTTGTGTGATGATGCAAAATGCTGCT 1140
QY      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
DB      1141 GATTCAAAGCTGCCAGAAATTAATGACTGGAAGAGATTTGAGGGGAAGTTTGTGG 1200
QY      401 AlaThrPheTYrProLeuSerAlaTYrLysArgGlyTYrLeuTYrGlnThrLeuLeu 419
DB      1201 GCTACATTTCTACCCGCTGATGCTCTTCAAGAGGGGATATGCTATCAAACTTGTCT 1257

```

RESULT 3
 AA166829 standard; cDNA, 1260 BP.
 ID AA166829
 XX AA166829;
 AC
 XX
 DT 07-JAN-2002 (first entry)
 XX
 XX Human protein kinase polypeptide 13887 coding sequence.
 DE
 XX Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
 KW cytostatic; antiinflammatory; immunosuppressive; cardiac; hepatotropic;

KM gene therapy; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1260
FT /tag= a
FT /product= "protein kinase 13887"
XX
XX MO200173050-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-US009483.
XX
XX 24-MAR-2000; 2000US-0191846P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers R;
XX
XX WPI: 2001-611632/70.
XX P-PSDB; AAG65767.
XX
XX
XX New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
PT useful in diagnosis of cancer or cellular proliferation or
PT differentiation disorders and to screen for polypeptide modulators useful
PT to treat such conditions.
XX
XX
XX Claim 1; Fig 10A-C; 169pp; English.
XX
XX The invention provides novel human protein kinase polypeptides, 3714,
CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC protein kinase polypeptides can be expressed by standard recombinant
CC methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC are useful for diagnostic and screening methods to identify subjects (at
CC risk of) having cancer or cellular proliferation and/or differentiation
CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC modulators are useful for the treatment of cancer, particularly colon
CC cancer or cellular proliferation and/or differentiation disorders. Other
CC disorders associated with 3714, 16742, 23546 or 13887 expression or
CC activity that can be treated include bone related disorders, inflammatory
CC disorders, autoimmune diseases, cardiovascular disorders and liver
CC diseases. The present sequence represents a human protein kinase
CC polypeptide 13887 coding sequence
XX
XX
SQ Sequence 1260 BP; 312 A; 279 C; 334 G; 335 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,75e-206 Length: 1260
Score: 2195.00 Matches: 417
Percent Similarity: 99.76% Conservative: 1
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 99.50% Indels: 0
DB: 4 Gaps: 0

US-10-798-532-4 (1-419) x AA166829 (1-1260)
QY 1 MetAlGlySerGlyCysAlaITPGLyAlaGluProProArghPheLeuGluAlaIaPheGly 20
Db 1 ATGGCGGGAATCCGAGCTGGCTGGGGGCGGAGCCGCGCTTTCTGGAGGGCTTCGGG 60
QY 21 ArgLeuITPGLInValGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArGVal 40
Db 61 CGGCTGTGGAGGATCAAGAGCCGCTGTGGTAGGGGCTCTCCGCTCGGTATCGGGTT 120
QY 41 ArgCysGlyGlnAlaPheProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
Db 121 CGCTGTGGCGGCAACCTGTGGCTGGCCCGCGGCGCTCTCAAGACAGTTCCTTGGCCGAGGA 180
QY 61 ThrThrGlyAlaIaIaIaSerAlaIaGluTYrGlyPheArgArgGlyAlaGlyAlaIaLeu 80
Db 181 ACCACCGGGGCTGGGCTGTGGCGCGGAGTATGTGTTCCGCAAAAGAGGGGCGGCTG 240

QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTYrGlyValPheThrIleHisPhe 100
Db 241 GAACAGTTCAGAGGTCAAGAAACATCGTGAATTTGTATGAGGTGTTAACTCACTTT 300
QY 101 SerProArghValProSerArgCysLeuLeuLeuGluLeuLeuAlaSerValSerGlu 120
Db 301 TCTCCAAAGTGCATCAAGCTGTGTGTGCTTGAACCTCGGAAGTCAAGTGTTCGGAA 360
QY 121 LeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTPMetIleGlnHisCysAlaArg 140
Db 361 TTGCTCTTATATTCAGTCAAGAGGCTTTCATGTGATGATACAGATTTGTCCGA 420
QY 141 AspValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTYrValHisAlaPheLeuYs 160
Db 421 GATGTTTGAAGCCCTTCTTTCTTCATCATGAGGGGTATGTCCATCGGAGCTCAAA 480
QY 161 ProArgAsnIleLeuITPSeSerAlaGluLeuGluCysPheLeuLeuIleAspPheGlyLeu 180
Db 481 CCACGTAAATATTTGGAGTGCAGAGATGAATGATGTTTAACTGATGACTTTGACTT 540
QY 181 SerPheLeuGluGlyAsnGlnAspValIleTYrIleGlnThrAspGlyTYrArgAlaPro 200
Db 541 AGCTTCAAGAGGCAATCAGATGTAAAGTATATTCAGACAGCGGATGCGGCTCA 600
QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
Db 601 GAAGCAGATTCCAAAATGCTTGGCCCGCGCTGCGACAGATGATCAAGATGTACC 660
QY 221 SerAlaValAlaPheLeuITPSeSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetYs 240
Db 661 TAGCTGTGAATCTGGAGGCTAGAGAAATCTTATCTGGAATGTTCAGAAATGAA 720
QY 241 LeuLYHisIThrValArgSerGlnGluTYrPlyAlaAsnSerSerAlaIleIleAspHis 260
Db 721 CTGAACATCAACAGTCAAGATCTCAGAAATGGAAGGCAACAGTTCGTATTTGATCAC 780
QY 261 IlePheAlaSerLYAlaValAlaValAlaIleProAlaTYrHisIleLeuArgPheLeu 280
Db 781 ATATTGGCAGTAAAGCAATGTGTAATGCCGAATTCACGCTTAACCTTAAGAGCTT 840
QY 281 IleLYSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
Db 841 ATCAAAAGATCTTCATATATATCCAAAGCAGAAAGATTCGTGAATGGCATTTGTC 900
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
Db 901 AGCCCATTTCTTATGCAATTCCTTTTCCCTCATATTTGAAGATCTGTCAATGCTCCACT 960
QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTYrLeuGlyAsnGluGluTYr 340
Db 961 CCAGTCTTAAGACTGCTGAATGTGCTGAGTATGATTTATCTTGAGAAATGAAAGGAAATAT 1020
QY 341 GluAspValValGluAspValLYSGluGluCysGlnLYSerGlyProValAlaSerLeu 360
Db 1021 GAAGATGTGTGAAGATGTAAAGAGAGGTCTCAAAATATGACCAAGTGTATCTCTCA 1080
QY 361 LeuValProLYSGluAsnProGlyArgGlyGlnValPheValGluTYrAlaAsnAlaGly 380
Db 1081 CTTGTTCCAAAGAAATCTCTGGCAGAGGACAAAGTCTTTGTGAGATGCAAAATGTGGT 1140
QY 381 AspSerLYAlaAlaGlnLYLeuLeuThrGlyArgMetPheAspGlyLYAspPheValVal 400
Db 1141 GATTCAAAGCTGCGCAGAAATTAAGTCACTGGAAGAGATTTGATGGGAAGTTGTTGTG 1200
QY 401 AlaThrPheTYrProLeuSerAlaTYrLYAsaArgGlyTYrLeuTYrGlnThrLeuLeu 419
Db 1201 GCTACATTTACCCGCTGTGTGCTTACAGAGGGGATATCTGTATCAAACTTGTCTT 1257
RESULT 4
ABQ93383
ID ABQ93383 standard; cDNA; 2008 BP.
XX

AC ABQ99383;
 XX 25-FEB-2003 (first entry)
 XX
 XX Human coding sequence SEQ ID 116.
 DE
 XX Human, expressed sequence tag; EST; haematopoietic disorder;
 KM central nervous system disease; viral infection;
 KM peripheral nervous system disease; non-healing wound; infectious disease;
 KM immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KM fungal infection; autoimmune disorder; coagulation disorder; neutrotropic;
 KM anti-allergic; antiinflammatory; immunosuppressive; neuroprotective;
 KM cytostatic; haemostatic; virucide; antibacterial; fungicide;
 KM immunostimulant; cerebroprotective; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200259260-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 16-NOV-2001; 2001MO-US042950.
 XX
 PR 17-NOV-2000; 2000US-00714936.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QX;
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 PI
 DR MPI; 2002-590824/63.
 XX N-PSDB; ABP64797.
 PT
 PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.
 XX
 PS Claim 1; SEQ ID NO 116; 394pp; English.
 XX
 CC The present invention relates to novel human coding sequences (ABQ99268-
 CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotentia or pluripotentia state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2008 BP; 491 A; 451 C; 488 G; 578 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,46e-206 Length: 2008
 Score: 2195.00 Matches: 417
 Percent Similarity: 99.76% Conservative: 1
 Best Local Similarity: 99.52% Mismatches: 1

Query Match: 99.50% Indels: 0
 DB: 6 Gaps: 0
 US-10-798-532-4 (1-419) x ABQ99383 (1-2008)
 QY 1 MetAlaGlySerGlyCysAlaATrpgAlaGluProPheArgPheLeuGluAlaPheGly 20
 Db ATGCGCGGATCCGGCTGCGCGCTGGGGCGCGAGCGCGCGCTTTTCGTGAGGCGCTTGGG 138
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerSerAlaSerValTyrArgVal 40
 Db CGGCTTGCGAGGTACAGAGCGCTGCGGTACGCGCTTCCGCGCTGTAATCGGGTT 198
 QY 41 ArgCysCysGlyAsnProGlySerProGlyAlaLeuArgGlnPheLeuProGly 60
 Db 199 CGTGCTGCGGCAACCTTGCTGCGCGCGCGCGCGCTTCAAGCAGTTCTTGCGCGCAGA 258
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
 Db 259 ACCACCGGGGCTGCGCGCTTGGCGCGAGTATGTTTCGCAAGAGAGGGCGGCGTG 318
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 Db 319 GAACAGTTCCAGGGTCCAGAAACATCGACTTGTATGAGTGTTCACATCCACTTT 378
 QY 101 SerProAsnValProSerArgCysLeuLeuGluLeuLeuAspValSerValSerGlu 120
 Db 379 TCTCCAAATGTGCATACGCTGTCTGTGTTGAATCTCTGAGATGCAAGTTCGAA 438
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 Db 439 TTGCTCTTATATTCAGTCACAGAGGGTGTTCATGTGATGATACAGCATTTGGCGCGA 498
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisIleGlnGlyTyrValHisAlaAspLeuLys 160
 Db 499 GATGTTTGGAGGCCCTTCTTCTTCATCATGAGGGTATGTCATCGACCTCAAA 558
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
 Db 559 CACGTAACATATTGTGGAGTGCAGAGATGATGTTTAACTCATTTGACTTTGGACTT 618
 QY 181 SerPheLeuGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
 Db 619 AGCTTAAAGAGGAGCAATCAGATGTAAAGTATTTACACAGACGGGTATGGGCTCCA 678
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 Db 679 GAAGCAGAAATTGCAAATTTGCTGGCCAGGCTGCGCTCAGAGTGATCAGAAATGTACC 738
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
 Db 739 TCGCTGTGATCTGGAGGCTTGAAGATCATTTTCTGAAATGTTCTCAGGAATGAAGA 798
 QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHis 260
 Db 799 CTGAACATATCAGTCACTCAGATGAAAGGCAACAGTTGCTATTTATGATCAC 858
 QY 261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
 Db 859 ATATTGGCCAGTAAGACATGTGGATGCGCAATTCACGCTTATCATCACTTAAGACCTT 918
 QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 Db 919 ATCAAAAGCATGCTTCATATGATTCACAGAGAAATTTCTGCTGAATGGCATTTGTC 978
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 Db 979 AGCCCATTTCTTAAGATCTTTTGGCCCTCATATTTGAAGATCTGGTATGCTTCCACT 1038
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTyrLeuGlyAsnGluGluGlyTyr 340
 Db 1039 CCAGTCTCAAGACTGCTGAATGTGCTGATGATGATATCTTGAGAAATGAAGAGAAATAT 1098

QY 341 GIUAAspValaIGluAspValIySGluGluCySGlnIySrgIyProValIaISerLeu 360
 DB 1099 GAAGAGTGTGTAGAAAGTGTAAAGAGAGTGTCAAAAATATGACACAGTGTATCTCTA 1158
 QY 361 LeuValProIySGluAspProGIyAGIyGlnValIyPheValIuTyraIaenIaAGIy 380
 DB 1159 CTGTGTCGAAAGGAAATCTGTGACAGAGCAAGTCTTGTATGATGCAATGTGTGGT 1218
 QY 381 AspSerIyAlaIaGlnIySLeuLeuThrGIyArgMetPheAspGIySLeuPheValIa 400
 DB 1219 GATTCGAAAGCTCGCGAATATCTGACTGGAAGATGTTATGGAAGTTGTGTG 1278
 QY 401 AlaThrPheTyProIeSerAlaTyIySArgGIyTyIeutyrgIinthIeuleu 419
 DB 1279 GCTACATTCTACCGCTGAGTGTCTACAGAGGGAGATCTGTATCAAACTGTGCTT 1335

RESULT 5
 AA166828
 ID AA166828 standard; cDNA; 2622 BP.
 AC AA166828;
 XX
 DT 07-JUN-2002 (first entry)
 DE Human protein kinase polypeptide 13887 encoding cDNA.
 XX
 KM Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
 KM cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatocytic;
 KM gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 5' UTR 1..266
 FT CDS /*tag= a
 FT 267..1529
 FT /*tag= b
 FT /product= "protein kinase 13887"
 FT /note= "coding sequence specifically claimed"
 FT 1530..2622
 FT 3' UTR /*tag= c

MO200173050-A2.
 04-OCT-2001.
 23-MAR-2001; 2001MO-US009483.
 XX
 PR 24-MAR-2000; 2000US-0191846P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Meyers R;
 XX
 DR WPI; 2001-611632/70.
 DR P-PSDB; AAG65767.
 XX
 PT New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,
 PT useful in diagnosis of cancer or cellular proliferation or
 PT differentiation disorders and to screen for polypeptide modulators useful
 PT to treat such conditions.
 PT
 PS Claim 1; Fig 10A-C; 169pp; English.
 XX
 CC The invention provides novel human protein kinase polypeptides, 3714,
 CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
 CC protein kinase polypeptides can be expressed by standard recombinant
 CC methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
 CC are useful for diagnostic and screening methods to identify subjects (at
 CC risk of) having cancer or cellular proliferation and/or differentiation
 CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
 CC modulators are useful for the treatment of cancer, particularly colon
 CC cancer or cellular proliferation and/or differentiation disorders. Other

CC disorders associated with 3714, 16742, 23546 or 13887 expression or
 CC activity that can be treated include bone related disorders, inflammatory
 CC disorders, autoimmune diseases, cardiovascular disorders and liver
 CC diseases. The present sequence represents a human protein kinase
 CC polypeptide 13887 encoding cDNA
 XX
 SQ Sequence 2622 BP; 652 A; 567 C; 627 G; 776 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5, 11e-206 Length: 2622
 Score: 2195.00 Matches: 417
 Percent Similarity: 99.76% Conservative: 1
 Best Local Similarity: 99.52% Mismatches: 1
 Query Match: 99.50% Indels: 0
 DB: 4 Gaps: 0

US-10-798-532-4 (1-419) x AA166828 (1-2622)

QY 1 MetAlaGIySerGIyCySAlaIATrGIyAlaGIuProIyArgPheLeuGIuAlaPheGIy 20
 DB 267 ATGCGCGGATCCGGCTGCGCTGCGGCGCGGCGCGCGCTTTCTGAGGCTTCGGG 326
 QY 21 ArgLeuTrpGIyValGlnSerArgLeuGIySerGIySerSerAlaSerValTyraArgVal 40
 DB 327 CGCTGTGCGAGATACAGGCCGTCTGAGTACGGCTCTCGCCTCGGTATCGGGTT 386
 QY 41 ArgCySGIyAsnProGIySerProProGIyAlaLeuArgGlnPheLeuProProGIy 60
 DB 387 CGCTGTGCGGACCCCTGTGCTGCGCCCGCGCGCTTCAACAGATTTGCGCGCAGGA 446
 QY 61 ThrThrGIyAlaAlaAlaSerAlaAlaGIyTyIyPheArgIyGluArgAlaAlaLeu 80
 DB 447 ACCACCGGGCTGCGGCTCTGCGCGAGTGTGTTCCGAAAGAGGCGCGCGCTG 506
 QY 81 GlnGlnLeuGlnGIyHisArgAsnIleValThrLeuTyGIyValPheThrIleHisPhe 100
 DB 507 GACAGTGTACAGAGTACAGAAACATCGTACTTGTATGAGTGTTCATTCACACTTT 566
 QY 101 SerProAsnValProSerArgCySLeuLeuGIyLeuLeuAspValSerValSerGIu 120
 DB 567 TCTCCAAATGTGCATACGCTGTGCTTGAACCTCTGAGTGTGAGTGTTCGGAA 626
 QY 121 LeuLeuLeuTySerSerHisGlnGIyCySerMetTrpMetIleGlnHisCySAlaArg 140
 DB 627 TTGCTCTTATATTCACATCACAGGAGTGTTCATGTGATGATACAGCATTTGCCCA 686
 QY 141 AspValLeuGIuAlaLeuAlaPheLeuHisGIyGIyTyValHisAlaAspLeuTyS 160
 DB 687 GATGTTTGGAGCCCTTGCTTTCTTCTCATGAGGCTATGTCCATCGGACCTCAAA 746
 QY 161 ProArgAsnIleLeuTrpSerAlaGlnGluGluCySLeuLeuIleAspPheGIyLeu 180
 DB 747 CCACGTAAATATTTGGAGTGTCCAGAGAAATGAATGTTTAAACTCATTCATTCGACTT 806
 QY 181 SerPheIyGluGIyAsnGlnAspValIyTyIleGlnThrAspGIyTyraArgAlaPro 200
 DB 807 AGCTTCAAAGAAAGGCAATCAGATGTAAAGTATATTCAGACAGAGCGGTATGGCTCCA 866
 QY 201 GIuAlaGlnLeuGlnAsnCySLeuAlaGlnAlaGIyLeuGlnSerAspThrGIyCySThr 220
 DB 867 GAAGCAGAAATGTCAAAATTTGCTTGCCAGGCTGGCTGCAAGTGTATCAGAAATGTAC 926
 QY 221 SerAlaValaAspLeuTrpSerLeuGIyIleIleLeuLeuGluMetPheSerGIyMetIyS 240
 DB 927 TCAGCTGTGATCTGTGAGGCTTAAAGATTTTAACTGGAATGTTCAGGAATGAAA 986
 QY 241 LeuIyHisIleThrValArgSerGlnGluTrpIyAlaAsnSerSerAlaIleIleAspHis 260
 DB 987 CTGAAGACATACAGTGTGATCTCAGAGATGGAAGGCAACAGTTCCTATTATTGATCAC 1046
 QY 261 IlePheAlaSerIyAlaAlaValaIaenIaIleProAlaTyHisIleuArgAspLeu 280
 DB 1047 ATATTGTCAGTAAAGCAAGTGTGTAATGCCGCAATTCCAGCTTATCATCAAGAGACCTT 1106

QY 281 ILeYSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 1107 ATCAAAAGCATGCTTCATGATGATCCAGCAGAGAGATTCCTGCTGAAATGCGATTGTGC 1166
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 1167 AGCCCATTTCTTATGATTCCTTTTGGCCCTCATATTGAAGATCTGGTCATGCTTCCCACT 1226
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspPheTyrIleuGluAsnGluGluTyr 340
 DB 1227 CCAAGTCTAAGATGCTGAAATGCTGATGATGATTTATCTTGAAGATGAAGAGAAATAT 1286
 QY 341 GluAspValValGluAspValValGluGluGluGluGluGluGluGluGluGluGluGlu 360
 DB 1287 GAAGATGTTGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGATGTAAGAT 1346
 QY 361 LeuValProGluGluAsnProGluArgGluGluValPheValGluTyrAlaAsnAlaGly 380
 DB 1347 CTGTTCCAAAGGAAATCTGGCAGAGCAGCAAGTCTTGTGATGATGCAAAATGCTGGT 1406
 QY 381 AspSerIleAlaAlaGluIleLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 DB 1407 GATTCCAAAGCTCGCAGAAATTAATGACTGGAAGAGATGTTGATGGAAGATTTGTGTG 1466
 QY 401 AlaThrPheTyrProLeuSerAlaTyrIleArgGlyTyrIleuTyrGlnThrLeuLeu 419
 DB 1467 GCTACATTCTACCCGCTGAGTGCCTACAGAGGAGATATCTGTAACACTTGTCTT 1523
 RESULT 6
 AAH72714
 ID AAH72714 standard; cDNA; 4065 BP.
 AC AAH72714;
 XX 19-SEP-2001 (first entry)
 XX Human cervical cancer marker nucleic acid 3988.
 DE Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
 XX Homo sapiens.
 XX MO200142467-A2.
 PN 14-JUN-2001.
 PD 08-DEC-2000; 2000MO-US033312.
 PF 08-DEC-1999; 99US-0169681P.
 PR 21-DEC-1999; 99US-0171350P.
 PR 14-MAR-2000; 2000US-0189315P.
 PR 12-MAY-2000; 2000US-0203791P.
 PR 09-JUN-2000; 2000US-0210600P.
 PR 21-JUL-2000; 2000US-0220114P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Deeds J, Berger A, Zhao X;
 PI WPI; 2001-375006/39.
 DR New isolated nucleic acid for diagnosing and treating cervical cancer and
 PT for assessing and detecting compounds for treating the cancer.
 XX Claim 1; Page 810-811; 1051p; English.
 XX The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful: to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for

CC Inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy
 XX
 SQ Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;
 Alignment Scores:
 Pred. No.: 9,696-206 Length: 4065
 Score: 2195.00 Matches: 417
 Percent Similarity: 99.76% Conservative: 1
 Best Local Similarity: 99.528 Mismatches: 1
 Query Match: 99.508 Indels: 0
 DB: 4 Gaps: 0
 US-10-798-532-4 (1-419) x AAH72714 (1-4065)
 QY 1 MetAlaGlySerGlyCysAlaIleTrrGlyAlaGluProProArgPheLeuGluAlaPheGly 20
 DB 268 ATGGCGGAGATCCGGCTGCGCCCTGGGGCGCGAGCCGCCGTTTCTGGAGGCTTTCGGG 327
 QY 21 ArgLeuTrrGluValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB 328 CCGCTTGGCAGATACAGAGCCCTGAGTACGGCTCTCCGCTCGATGATCGGTT 387
 QY 41 ArgCysCysGlyAsnProGlySerProProGluAlaLeuArgGlnPheLeuProGly 60
 DB 388 CGCTGCTGGGCAACCTGCTCGCCCGCGCGCCCTCAACAGATTTCTTCCCGCAGGA 447
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgLysGluArgAlaAlaLeu 80
 DB 448 ACCACCGGGGCTGGCGCTCTGCTCGCCCGCGATGATTTCCGAAAGAGGGCGCGCTG 507
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB 508 GAAACGTTCAAGGCTCACAGAAACATGTAATCTTGATGAGTGTTCATTCACATCTT 567
 QY 101 SerProAsnValProSerArgCysLeuLeuGluGluLeuAspValSerValSerGlu 120
 DB 568 TTCGCAATGTGCATCAACGCTGTGCTTGAACCTCTGATGATGATGATGATGATGATGAT 627
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrrMetIleGlnHisCysAlaArg 140
 DB 628 TTGCTCTTATATTCAGTCCAGCAGAGGTTGTTCCATGTGATATCAGATGATGATGATGAT 687
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisGluGluTyrValHisAlaAspLeuVal 160
 DB 688 GATGTTTGAAGCCCTTGTCTTCTCATGATGAGGCTATGATCCATGCGAATCTCAAA 747
 QY 161 ProArgAsnIleLeuTrrPheSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu 180
 DB 748 CCACTGAATATTTGAGAGTGCAGAGATGATGATGATGATGATGATGATGATGATGATGAT 807
 QY 181 SerPheLysGluGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
 DB 808 AGCTTCAAAAGAGGCAATCAGATGTAAGTATATTCAGACAGAGGCTATGCGGCTCCA 867
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 DB 868 GAAGCAGAAATGCAAAATGCTTGGCCAGGCTGCGCAGAGATGATCAAGAAATGATCC 927
 QY 221 SerAlaValAspLeuTrrPheSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLys 240
 DB 928 TCAGCTGTTGATCTGTGGAGCTTGAAGATCAATTTAACTGGAATGTTCTCAGAGATGAA 987
 QY 241 LeuLysHisThrValArgSerGlnGluTrrLysAlaAsnSerSerAlaIleIleAspHis 260
 DB 988 CTGAACAATACAGTCAATCTCAGAAATGGAAGGCAACAGTTCTGATATTTATATCAC 1047
 QY 261 IlePheAlaSerLysAlaAlaValAlaAsnAlaAlaIleProAlaTyrHisIleuAspLeu 280
 DB 1048 ATATTGTCAGTAAGAGATGTAATGCGCAATTCACGCTTATCACTTAAGAACCTT 1107
 QY 281 ILeYSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300

Db 1108 ATCAAAAGCATGCTTCATGATGATCCAGCAGAAAGATTCTGCTGAATGCGATTGTC 1167
Qy SerProPhePheSerIleProPheAlaProHisIleGluIleuPheValMetIleuProThr 320
Cc 301 ATGAGGATGCTTTTGGATCTTTTGGCCCTCATATTGAGATCTGGTCATGCTTCCCACT 1227
Db 1168 AGCCCATCTTTTGGATCTTTTGGCCCTCATATTGAGATCTGGTCATGCTTCCCACT 1227
Qy 321 ProValIleuArgIleuIleuValIleuAspAspAspTyrIleuGluIleuGluIleuTyr 340
Cc 1228 CCACTGCTAAGCTGCTGAAATGCTGAGATGATTAATCTTGAGAAATGAGAGAAATAT 1287
Db 341 GluAspValValIleuAspValIleuGluIleuGluIleuTyrIleuProValIleuSerIleu 360
Cc 1288 GAAAGATGTTTGAAGATGTTTGAAGAGAGAGATGTCAAAATATATGACCGATCTCTA 1347
Qy 361 LeuValIleuArgIleuIleuProGluIleuGluIleuValIleuPheValIleuTyrIleuValIleuAsnIleu 380
Cc 1348 CTGCTTCCAAAGAGAAATCTTGGCAGAGACCAAGTCTTTTGAAGATGCAATGCTGCT 1407
Db 381 AspSerIleuAlaIleuGluIleuIleuIleuTyrGluIleuPheAspGluIleuPheValIleu 400
Cc 1408 GATTCCAAAGCTGCGCAGAAATTAATGACTGAAAGATGTTGATGGAAGTGTGTTG 1467
Qy 401 AlaThrPheThrProIleuSerAlaTyrIleuArgIleuTyrIleuTyrIleuIleu 419
Cc 1468 GCTACATTCATCCCGCTGAGTGCCTACAGAGGAGATATCTGATCAAACTTGCTT 1524
Db

RESULT 7

ADL45238
ID ADL45238 standard; DNA; 4065 BP.

AC ADL45238;

DT 20-MAY-2004 (first entry)

XX Human ovarian cancer DNA marker #19128.

XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.

XX Homo sapiens.

XX WO200170979-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009126.

XX 21-MAR-2000; 2000US-0191031P.

XX 15-JUN-2000; 2000US-0207124P.

XX 07-JUL-2000; 2000US-0211940P.

XX 25-JUL-2000; 2000US-0216820P.

XX 21-DEC-2000; 2000US-0257672P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lee J, Lillie J;

XX WPI; 2001-611502/70.

XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian cancer cells as compared to their normal non-cancerous ovarian cells are used to characterize stage, grade, histological type of ovarian cancer.

XX Disclosure; SEQ ID NO 19128; 106pp; English.

XX The invention relates to nucleic acid markers which are overexpressed in ovarian cancer cells as compared to their expression in normal (i.e. non-cancerous) ovarian cells. The invention also relates to polypeptides encoded by the markers, antibodies that selectively bind to the polypeptides, a method of inhibiting ovarian cancer in a patient at risk of developing ovarian cancer involving inhibiting expression of a gene corresponding to a marker of the invention and a method of treating a patient afflicted with ovarian cancer comprising providing to cells of

Cc the patient an antisense oligonucleotide complementary to a marker of the Cc invention. The markers are useful for assessing if a patient is afflicted Cc with ovarian cancer, which involves comparing the level of expression of Cc a marker in a patient sample and a normal level of expression of the Cc marker in a control non-ovarian cancer sample. A difference between the Cc expression levels indicates ovarian cancer. The level of expression of a Cc marker corresponds to a secreted protein or to a transcribed Cc polynucleotide or its portion. The level of expression of the marker is Cc assessed by detecting the presence in the sample, a protein or protein Cc fragment corresponding to the marker. The presence of protein or protein Cc fragment is detected using an antibody that specifically binds with the Cc protein or protein fragment. Alternatively, the level of expression of Cc the marker is assessed by detecting the presence of a transcribed Cc polynucleotide which anneals with the marker or anneals with a portion of Cc the polynucleotide comprising the marker, under stringent conditions. The Cc marker is also used for monitoring the progression of ovarian cancer in a Cc patient which involves detecting expression of the marker in a patient Cc sample at a first point in time, repeating the method at a subsequent Cc time and comparing the level of expression. The method is carried out Cc using an ovarian tissue sample. A composition comprising a marker, Cc polypeptide or antibody of the invention is used to treat ovarian cancer. Cc This sequence represents a human ovarian cancer DNA marker of the Cc invention.

XX Sequence 4065 BP; 1068 A; 822 C; 905 G; 1265 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.: 9,69e-206 Length: 4065
Score: 2195.00 Matches: 417
Percent Similarity: 99.76% Conservative: 1
Best Local Similarity: 99.52% Mismatch: 1
Query Match: 99.50% Indels: 0
DB: Gaps: 5

US-10-798-532-4 (1-419) x ADL45238 (1-4065)

Qy 1 MetAlaGlySerGlyCysAlaTrrpGluValIleuPheProPheIleuGluValIleuGly 20
Db 268 ATGGCGGGATCCGGCTGCGCCCTGGGGCGCGGCGCGCGCTTTCTGAGGCTTCGGG 327
Qy 21 ArgLeuTrrpGluValIleuSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db 328 CCGCTTGGCAGGTACAGAGCCCTCTGGGTAGCGGCTCTCGCTCGGTGTATGCGGT 387
Qy 41 ArgCysCysGlyAsnProGlySerProProGluAlaIleuArgGlnPheIleuProGly 60
Db 388 CGCTGCTGGCAACCTGCTGCGCCCGCGCCCTCAACAGATCTTTCGCGCAGGA 447
Qy 61 ThrThrGlyAlaAlaAlaSerAlaIleuTyrGlyPheArgGlyGluArgAlaIleu 80
Db 448 ACCACCGGGGCTGGCGCTCTGCGCCGCGAGTATGATTTCCGAAAGAGAGGCGCGCTG 507
Qy 81 GluGlnIleuGlnGlyHisArgAsnIleValThrLeuTyrGlyValIleuPheThrIleuPhe 100
Db 508 GAACAGTTGACGGTTCACAGAAACATCGTGAATTTGTATGAGTGTATTAATCACTTT 567
Qy 101 SerProAsnValProSerArgCysLeuIleuGlnIleuAspValSerValSerGlu 120
Db 568 TCTCCAAATGTCATCAACGCTGTCTGTGTAACCTCTGAGATGATGATGATGATG 627
Qy 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrrpMetIleuGlnHisArgValArg 140
Db 628 TTGCTCTTATATTCAGTCAACAGAGGTGTTCATGATGATATACAGATTTGTCGGA 687
Qy 141 AspValIleuGluAlaIleuAlaPheIleuHisGlnGlyTyrValIleuAlaIleuPheIleu 160
Db 688 GATGTTTGGAGGCCCTTCTTTTCTTCAATCAATGAGGCTATGTCATGCGGACCTCAA 747
Qy 161 ProArgAsnIleuLeuTrrpSerAlaGluAsnGluCysPheValIleuIleuAspPheIleu 180
Db 748 CCACGTAACATATTTGGAGGTGACAGAGATGAATGTTTAACTCATTCATTCGACTT 807
Qy 181 SerPheIleuGluGlyAsnGlnAspValIleuTyrIleuGlnIleuAspGlyTyrArgAlaPro 200


```

Db      808 AGCTCAAGAGGACATGATGATTAATATATTCAGACAGCGGATATGGCTCCA 867
Qy      201 GUAAGAGUleuGlnAaenCySleuAlaGlnAlaGlyLeuGlnSerPheThr 220
Db      868 GAAGCGAATTGGAAAATGCTTGGCCCGAGCTGGCTGCAAGATATACAGATATACC 927
Qy      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuGlnLuePheSerGlyMetLys 240
Db      928 TCAAGCTTGTATGTTGAGAGCTTGAAGATTCATTTTCTGGAAATGTTCTCAGAAATGAAA 987
Qy      241 LeuLysHisThrValArgSerGlnIuTrpLysAlaAsnSerSerAlaIleIleAspHis 260
Db      988 CTGAAACATATACATGATCTGATCTGAGAAATGAGAAAGCAAGTTCTGCTATATATGATCAC 1047
Qy      261 IlePheAlaSerLysAlaValAlaAlaAlaIleProAlaTyrtHisLeuAlaPheLeu 280
Db      1048 AATATTGGCCAGTAAACAGAGTGGTGAATGGCCGAATTCACAGCTTATCAGACAGACTTT 1107
Qy      281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuGly 300
Db      1108 ATCAAAAGCATGCTTATGATGATCCAGACGAAAGAAATTCCTGCTGAATGGCATTTGTC 1167
Qy      301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db      1168 AGCCCATTTCTTATGATCTCTTTTGGCCCTCATATGAAATGATGCTCATGCTCCCACT 1227
Qy      321 ProValLeuArgLeuLeuAlaValLeuAspAspAspTyrtLeuGlnValGlnGlnGlnGly 340
Db      1228 CCAGTCTCAAGATCTGCTGAATGCTGATGATGATTAATCTTGAATGAATGAAGAGAAATAT 1287
Qy      341 GlnAspValValGlnAspValLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
Db      1288 GAAGATGTTGTAAGAAATGTAAGAAAGAGTGTCAAAATATGACACAGGATCTCTA 1347
Qy      361 LeuValProLysGlnAsnProGlyArgGlnValPheValGlnTyrtAlaAlaAlaGly 380
Db      1348 CTTGTTCCAAAGAAATCTCTGCAAGACAGCAAGTCTTTGTCAGATGCAATGCTGCT 1407
Qy      381 AspSerLysAlaAlaGlnLysLeuLeuTrpGlyValGlnMetPheAspGlyLysPheVal 400
Db      1408 GATTCCAAAGCTGCGCAAAATTAATCACTGCAAGAGATGTTGATGGAAGTTTGTGTG 1467
Qy      401 AlaThrPheTyrtProLeuSerAlaTyrtLysArgGlyTyrtLeuTyrtGlnThrLeuLeu 419
Db      1468 GCTACATCTACCCGCTGAGTGTCTCAAGAGGGGATATCTGTATCAAACTTGTCTT 1524

```

RESULT 8
AAS06734
ID AAS06734 standard; cDNA; 1260 BP.

AC AAS06734;
DT 12-SEP-2001 (first entry)
XX
XX
DE Polynucleotide sequence encoding human protein kinase #34.
XX
XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disease;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX
XX Homo sapiens.
OS
PN WO200138503-A2.
PD 31-MAY-2001.
XX
XX 22-NOV-2000; 2000MO-US032085.
PR 24-NOV-1999; 99US-0167482P.
XX
PA (SUGEN-) SUGEN INC.

```

XX      XX Plowman GD, Whyte D, Manning G, Sudareanam S, Martinez R;
PI      PI Flanagan P, Clary D;
XX      DR WPI; 2001-343950/36.
XX      DR P-PSDB; AAO03534.
XX      PT Nucleic acids encoding human kinase polypeptides, useful for preventing
PT      diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT      neuronal-associated diseases, and microbial infections.
XX      XX
XX      PS Example 1; Fig 1; 433p; English.
XX      XX
CC      AAS06701-AAS06757 encode for novel human protein kinases #1-57. The novel
CC      protein kinases have been identified as members of the tyrosine or
CC      serine/threonine kinase (PTK and STK) families. The polynucleotides
CC      encoding protein kinases and the polypeptides may be used in the
CC      prevention, diagnosis and treatment of diseases associated with
CC      inappropriate kinase expression. For example, they may be used to treat
CC      cancers (especially cancers of haematopoietic origin), cardiovascular
CC      disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC      immune related diseases (e.g. rheumatoid arthritis), neurological
CC      disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC      Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC      disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC      Additionally, polynucleotides encoding protein kinases may be used for
CC      gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC      polypeptides may be used as antigens in the production of antibodies
CC      against the protein kinases and in assays to identify modulators of
CC      protein kinase expression and activity
XX      XX
SQ      Sequence 1260 BP; 311 A; 279 C; 335 G; 335 T; 0 U; 0 Other;

```

Alignment Scores:

| Pred. No.: | 8-59e-206 | Length: | 1260 |
|------------------------|-----------|---------------|------|
| Score: | 2188.00 | Matches: | 416 |
| Percent Similarity: | 99.52% | Conservative: | 1 |
| Best Local Similarity: | 99.28% | Mismatches: | 2 |
| Query Match: | 99.18% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-10-798-532-4 (1-419) x AAS06734 (1-1260)

```

Qy      1 MetAlaGlySerGlyCysAlaIleTrpGlyAlaGlnProProArgPheLeuGlnAlaPheGly 20
Db      1 ATGGCGGATTCGGCTGCGCTGGGAGCGCGCGGCTTCTGGAGGCTTCGGG 60
Qy      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrtArgVal 40
Db      61 CGGCTGTGGCAGATACAGAGCCGTCTGGGTAGCGGCTCTCGGCTCGGTATGGGTT 120
Qy      41 ArgCysArgLysAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
Db      121 CGCTGCTGGCGCAACCTGGCTGCGCCCGCGGCTTCAACAGATTTCTTGGCCAGGA 180
Qy      61 ThrTrpGlyAlaAlaAlaSerAlaIleGlnTyrtGlyPheArgGlyGlnAlaGlnAlaLeu 80
Db      181 ACCACCGGGGCTGGCGCTGCGCCCGCAGATATGATTTCCGAAAGAGGGCGGCGCTG 240
Qy      81 GlnGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrtGlyValPheThrIleHisPhe 100
Db      241 GAACAGTTCAGAGGTCACAGAAACATCGTGAATTTGATGAGATGTTTCAATCCACTT 300
Qy      101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGln 120
Db      301 TCTCCAAATGTGCATCAAGCTGTCTGTGAATCTCTGATGATGATGATGATGATGATG 360
Qy      121 LeuLeuLeuTyrtSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisGlyValArg 140
Db      361 TTGCTTTATATATTCAGTACAGAGGATGTTTCAATGATGATGATGATGATGATGATG 420
Qy      141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlnGlyTyrtValHisAlaAspLeuLys 160

```

```

Db      421 GATGTTTGGAGGCCCTTGCTTTCTTCATCATGAGGCGTATGTCATGCGGACCTCAAA 480
QY      161 ProArgAsn11LeuTrpSer1aGluAenGluCysPheIysLeu11LeaPheGlyLeu 180
Db      481 CCAAGCAACATATTGAGAGTGCAGAGATGATGTTTAACTCATTTGACTTGGACCTT 540
QY      181 SerPheIysGluGlyAenGlnAenValIysTyr11LeuThrAapGlyTyrAraGalaPro 200
Db      541 AGCTTCAGAGAGGACATGAGATGTAAGTATATTCAGACAGACGGGATCGGGCTCCA 600
QY      201 G1uAlaGluLeuGlnAenCysLeuAlaGlnAlaGlyLeuGlnSerAerPhrGluCysThr 220
Db      601 GAAGCAGAAATTGCAAAATGCTTGCGCCAGCGCTGCAAGATACAGATATGATACC 660
QY      221 SerAlaValAapLeuTrpSerLeuGlyIle11LeuLeuGlnUerPheSerGlyMeLys 240
Db      661 TCAGCTGTTGATCTTGAGACCTTACGAAATCTTTTACTGAAATGTTCTCAGAAATGAAA 720
QY      241 LeuLysIsthrValAArgSerGlnGluTrpLysAlaAenSerSerAlaIle11LeaPheHis 260
Db      721 CTGAAACATPACGATCTCAGAAATGGAAGCAAAACAGTTCTGCTATTATTGATCAC 780
QY      261 11ePheAlaSerIysAlaValAlaAenAla11eProAlaTyrHisLeuAraAapLeu 280
Db      781 ATATTGCCCAGTAACAGATGTAATGCCGCAATTCACGCTTATCACTTAAGACACTT 840
QY      281 11eLysSerMetLeuHisAapPProSerAraArg11eProAlaGluMeLalaLeuCys 300
Db      841 ATCAAAAGCATGCTTCATGATGATCCAGACAGAAATTCCTCTGTAATGCGATTGTGC 900
QY      301 SerProPhePheSer11eProPheAlaProHis11eGluAapLeuValMeLysProThr 320
Db      901 AGCCCATCTTTAGCATTCCTTTTGCCCTCATATTGAAGATCGGTCATGCTTCCACT 960
QY      321 ProValLeuAraGluLeuAenVal11eAapAapPProTyrLeuGlnGluGluTyr 340
Db      961 CCAAGTCTAAGACTGCTGAATGTGCTGATGATGATTAATCTTGAAGATGAAGGAAATAT 1020
QY      341 GluAapValAlaGluAapValIysGluGluCysGlnIsthrTyrGlyProVal11eSerLeu 360
Db      1021 GAAGATCTTCTTAAGATGTAATGAAGAGACTGTCAAAATAATATGACAGCGATCTCTA 1080
QY      361 LeuValProIysGluAenProGlyAraGlyGlnValA1pHeVal11uTyrAlaAenAlaGly 380
Db      1081 CTGTGTCAAAGGGAATCTGCGACAGAGCAAGTCTTTTGTAGTATGCAATATGCTGCT 1140
QY      381 AapSerLeuAla11aGlnIsthrLeuThrGlyAraGwerPheAapGlyLysPheVal11a 400
Db      1141 GATTCCAAAGCTCGCGAAATTAAGTGAAGGATGTTGATGGGAAATTTGTGTG 1200
QY      401 AlaThrPheTyrProLeuSerAlaTyrLysAraGlyTyrLeuTyrGlnThrLeuLeu 419
Db      1201 GCTACATTTCAACCCGCTGAGTGCCTACAAAGGGGATATCTGATCAAACTTGCTT 1257

RESULT 9
AADI8818
ID      AADI8818 standard; cDNA; 1824 BP.
XX
AC      AADI8818;
XX
DT      18-DEC-2001 (first entry)
XX
DE      Human Kinase (PKIN) -3 CDNA.
XX
KW      Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;
myaesthesia gravis; cirrhosis; cataract; growth and development disorder;
seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
asthma; obesity; restorative therapy; cytostatic; immunomodulatory;
antimicrobial; cardiovascular; antiinflammatory; vaccine; ss.

```

```

XX      OS      Homo sapiens.
XX      FH      Key      Location/Qualifiers
XX      FT      CDS      169..1428
XX      FT      /tag= a
XX      FT      /product= "Human PKIN-3 protein"
XX      PN      MO200181555-A2.
XX
XX      PD      01-NOV-2001.
XX
XX      PF      20-APR-2001; 2001W0-US012992.
XX
XX      PR      20-APR-2000; 2000US-0199021P.
XX      PR      28-APR-2000; 2000US-0200226P.
XX      PR      05-MAY-2000; 2000US-0202339P.
XX      PR      11-MAY-2000; 2000US-0203505P.
XX      PR      18-MAY-2000; 2000US-0205564P.
XX      PR      26-MAY-2000; 2000US-0207739P.
XX      PR      01-JUN-2000; 2000US-0208795P.
XX
XX      PA      (INCYTE GENOMICS INC.
XX
XX      PI      Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
PI      Bandman O, Lu DM, Lal P, Burford N, Khan FA, Walla NK, Yao MG;
PI      Patterson C, Burrill JD, Marcus GA, Zingler KA, Reclipon SA, Lu Y;
PI      Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
PI      Walsh RT, Rankumar J, Borowsky ML, Au-Young J, Hillman JL;
PI      Gururajan R;
XX
XX      DR      WPI; 2001-611740/70.
XX      DR      P-PSDB; AAB11769.
XX
XX      PT      Human kinases and nucleic acids, useful for preventing diagnosing and
XX      PT      treating cancers, inflammation and immune disorders.
XX
XX      PS      Claim 5; Page 152; 166pp; English.
XX
XX      CC      The present invention relates to human kinases (PKIN) and the nucleic
XX      CC      acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
XX      CC      used in the prevention, diagnosis and treatment of diseases cancers,
XX      CC      adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
XX      CC      acquired immune deficiency syndrome (AIDS), anaemia, asthma, allergies,
XX      CC      gout, microbial infections, cardiovascular disease and/or inflammation,
XX      CC      myaesthesia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
XX      CC      infarction, cataract, growth and development disorder, seizure disorder,
XX      CC      pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
XX      CC      disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
XX      CC      PKIN may be used to treat disorders associated with decreased PKIN
XX      CC      expression by rectifying mutations or deletions in a patient's genome
XX      CC      that affect the activity of PKIN by expressing inactive proteins or to
XX      CC      supplement the patients own production of PKIN. PKIN nucleic acids may be
XX      CC      used to produce the PKIN polypeptide, by inserting the nucleic acids into
XX      CC      a host cell and culturing the cell to express the protein. PKIN nucleic
XX      CC      acid and its complementary sequences may also be used as DNA probes in
XX      CC      diagnostic assays to detect and quantitate the presence of similar
XX      CC      nucleic acid sequences in samples and therefore which patients may be in
XX      CC      need of restorative therapy. The present sequence is human PKIN-3 CDNA
XX
XX      SQ      Sequence 1824 BP; 434 A; 416 C; 472 G; 502 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1.47e-205
Score:          2188.00      Length:      1824
Percent Similarity: 99.52%      Matches:      416
Best Local Similarity: 99.28%      Mismatches: 1
Query Match:      99.18%      Indels:      2
DB:              Gaps:      0

US-10-798-532-4 (1-419) x AADI8818 (1-1824)
QY      1 MeAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20

```

```

Db      169 ATGGCGGATCCGGCTGGCGGCGGAGACCGCGGCTTTCTGGAGGCGCTCGGG 228
Qy      21 ArgLeuTTGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      229 CGGCTGTGGCAGGTACAGAGCCGCTGGGTAGGGGCTCCGCCCTCGGGTATCGGGTT 288
Qy      41 ArgCysCysGlyAsnProGlySerProGlyValAlaLeuArgGlnPheLeuProProGly 60
Db      289 CGGTGCTGGGCAACCTGGCTGGCCCCCGGCGCCCTCAAGCATTTCTTGGCCGACGA 348
Qy      61 ThrThGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
Db      349 ACCACCGGGGCTCGGCGCTTGGCGCGGATGATGTTTCCGCAAGAGGGGGCGCTG 408
Qy      81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db      409 GAACAGTTTCAGGGTACAGAAACATCGTGACTTTGTATGAGGTGTTTCAATCCACTTT 468
Qy      101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuAspValSerValSerGlu 120
Db      469 TCTCAAAATGTCACACGCTCTGTGCTTGAACCTCTGAGATGCACTGTTTGGAA 528
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db      529 TTGCTCTTATATTCAGTACACAGGGTGTTCATGTGATGATACAGCATTTGCCCCA 588
Qy      141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGluGlyTyrValHisAlaAspLeuLys 160
Db      589 GATGTTTGGAGGCCCTTGCTTTCATCATGAGGGCATATCCATGCGGACCTCAAA 648
Qy      161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLeuLeuLeuAspPheGlyLeu 180
Db      649 CACGTAACATATTTGAGTGCAGTGCAGAAATGATTTTAACTCATTTGACTTGGACTT 708
Qy      181 SerPheLeuGlnGlyAsnGlnAspValLysTyrIleGlnThrAspGlyTyrArgAlaPro 200
Db      709 AGCTTCAAAAGAGGCAATCAGATGTAATTAATTCAGACAGACGGGTATCGGGCTCA 768
Qy      201 GluAlaGluLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
Db      769 GAGCGAGAAATGCAAAATGCTTGGCCCGAGGCTGAGAGTATCAGATATGATATACC 828
Qy      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnLeuPheSerGlyMetLys 240
Db      829 TACGCTGTGATCTGGAGCTTGAAGATATTTACTGAAATGTTCTCAGAAAGAA 888
Qy      241 LeuLysHisThrValArgSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHis 260
Db      889 CTGAACATACAGTCAAGATCTCAGAAATGAGAAAGCAAGCTTCTGCTATTAATGATCAC 948
Qy      261 IlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTyrHisLeuArgAspLeu 280
Db      949 ATATTTTCCAGTAAACAGTGGTGAATGCGCAATTCACGCTTATCACTTAAGACCTT 1008
Qy      281 IleLysSerMetLeuHisAspAspProSerArgIleProAlaGlnMetAlaLeuCys 300
Db      1009 ATCAAAAGCATGCTTATGATGATCCAGAGCAAGAAATTCCTGCTGAATGGCATTTGTC 1068
Qy      301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
Db      1069 AGCCCATTTCTTATGATCTCTTTGGCCCTCATATTTGAAGATGTGGTCAATGCTCCACT 1128
Qy      321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlnAsnGluGlnGluTyr 340
Db      1129 CCAGTCTTAAGATGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1188
Qy      341 GluAspValValGluAspValLysGluGlnCysGlnLysTyrGlyProValIleSerLeu 360
Db      1189 GAAGAGTGTGTAAAGATGTAAAGAGAGAGTGTCAAAATATGACACAGGGTATCTCA 1248
Qy      361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380

```

```

Db      1249 CTTGTTCCAAAGGAAATCTTGCGAGAGACAGTCTTTGTGATGATCAAAATCTGCT 1308
Qy      381 AspSerValAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db      1309 GATTCAAAGCTGCGCAGAAATTAATCTGAGAGAGATGTTGATGGAGATTTGTGG 1368
Qy      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db      1369 GCTACATTTACCCCTGAGTGCCTACAAAGAGGATATCTGTATCAAACTTGCTT 1425

RESULT 10
ADBS2857
ID      ADBS2857 standard; DNA; 3244 BP.
XX
AC      ADBS2857;
XX
DT      04-DEC-2003 (first entry)
XX
DE      Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3399.
XX
KW      toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW      toxicity marker; toxicity progression; drug screening;
KW      primary rat hepatocyte toxicity modelling; gene; db.
XX
OS      Rattus norvegicus.
XX
PN      W02003065993-42.
XX
PD      14-AUG-2003.
XX
PF      04-FEB-2003; 2003W0-US003482.
XX
PR      04-FEB-2002; 2002US-0353171P.
PR      13-MAR-2002; 2002US-0363534P.
PR      08-APR-2002; 2002US-0370248P.
PR      10-APR-2002; 2002US-0371134P.
PR      10-APR-2002; 2002US-0371135P.
PR      10-APR-2002; 2002US-0371150P.
PR      11-APR-2002; 2002US-0371413P.
PR      19-APR-2002; 2002US-0373601P.
PR      19-APR-2002; 2002US-0373602P.
PR      22-APR-2002; 2002US-0374139P.
PR      08-MAY-2002; 2002US-0378370P.
PR      09-MAY-2002; 2002US-0378652P.
PR      09-MAY-2002; 2002US-0378653P.
PR      09-MAY-2002; 2002US-0378655P.
PR      09-JUL-2002; 2002US-0394230P.
PR      09-JUL-2002; 2002US-0394253P.
PR      04-SEP-2002; 2002US-0407688P.
PR      28-JAN-2003; 2003US-0442900P.
XX
PA      (GENE-) GENE LOGIC INC.
XX
PI      Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI      Elashoff M;
XX
DR      WPI; 2003-731472/69.
XX
PT      Determining if a compound induces a toxic effect on a tissue or cell, for
PT      identifying hepatotoxic compounds, comprises comparing a gene expression
PT      profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT      mean values.
XX
PS      Claim 44; SEQ ID NO 3399; 874bp; English.
XX
CC      The present invention describes a method for determining whether a
CC      compound induces a toxic effect on a tissue or cell. The method comprises
CC      preparing a gene expression profile of a tissue or cell sample exposed to
CC      the compound, and comparing the gene expression profile to a database
CC      comprising data or information on the Tox mean and non-Tox mean value.
CC      The method is useful for predicting or identifying at least one toxic
CC      effect, particularly hepatotoxicity, of a test or unknown compound. The
CC      genes listed in the specification are useful as diagnostic or toxicity

```

CC markers for the prediction or identification of the physiological state
 CC of tissue or cell sample that has been exposed to a compound, or to
 CC identify or predict the toxic effects of a compound or an agent. These
 CC may also be used as markers for monitoring toxicity progression or for
 CC drug screening. The present sequence represents a primary rat hepatocyte
 CC toxicity modelling related gene sequence from the present invention.

XX Sequence 3244 BP; 800 A; 710 C; 819 G; 915 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 1, 07e-204 | Length: | 3244 |
|------------------------|------------|---------------|------|
| Score: | 2183.00 | Matches: | 414 |
| Percent Similarity: | 99.52% | Conservative: | 3 |
| Best Local Similarity: | 98.81% | Mismatches: | 2 |
| Query Match: | 98.96% | Indels: | 0 |
| DB: | 10 | Gaps: | 0 |

US-10-798-532-4 (1-419) x ADB52857 (1-3244)

```

QY 1 MetAlaGlySerGlyCybaLaATPGIyAlaGluProProAlaRPhLeuGluAlaPheGly 20
DB 162 ATGGCCGGGTCGGCTGGCGGTGGGGGCGGAGCCGCGCTTCTCGAGGCGCTTCGGG 221
QY 21 ArgLeuITrGlnAlaGlnSerArgLeuGlySerSerAlaSerValTYrArgVal 40
DB 222 CGGCTGTGGAGGTACAGAGCCGCTAAGGCGGCGCTCCCGGCTTCGGGTACCGGGTG 281
QY 41 ArgCybaGlyAlaSerProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
DB 282 CGCTGCTGGCGGACTCAAGGCTGCGCCCGCGGCGCTCAAGAGATTCCTGCTCCGGGA 341
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaLeu 80
DB 342 ACCACCGGGGCTCCGCTCGGCTGGCGAGTATGTTCCGAAAGAGAGGCGAGCGCTG 401
QY 81 GluGlnLeuGlnGlnHisArgAlaIleValThrLeuTYrGlyAlaPheThrIleHisPhe 100
DB 402 GAGCAGTGTGAGGGTACAGAGAACATCGACTTTTACGAGATCTTTACCATACACTTC 461
QY 101 SerProAlaValProSerArgCybaLeuLeuGluLeuLeuAlaSerValSerGlu 120
DB 462 TCTCCAAATGTGCATCACGCTCTGTTGCTTGAAGTCTTGAGATGACGCTTGGAA 521
QY 121 LeuLeuLeuTYrSerSerHisGlnGlyCybaSerMetTrpMetIleGlnHisCybaAlaArg 140
DB 522 TTGCTGTGATATTCAGTCCAGGAGGCTCATGATGATGCCAGACTGTGCCAG 581
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTYrTYrAlaHisAlaAspLeuLys 160
DB 582 GATGTCTGTGAGGCGCTTGTCTTTTTCACATGAGGCTATGTCCATGACAGCTCAAG 641
QY 161 ProArgAsnIleLeuTrpSerAlaGluAlaGluCybaPheLeuLeuAspPheGlyLeu 180
DB 642 CCACGAAACATCTCTGTGAGTGGCGGAGAACGATGCTTTAAGCTTTATGACTTGAATC 701
QY 181 SerPheLeuGlnGluAlaGlnAlaAspValLysTYrIleGlnThrAspGlyTYrArgAlaPro 200
DB 702 AGCTTCAAGAGAGCATCAGACGTGAAGTATATTCAGACAGAGGATATCAGACTCTCT 761
QY 201 GluAlaGlnLeuGlnAlaGlnAlaGlnAlaGlnAlaGlnAlaGlnAlaGlnAlaGln 220
DB 762 GAAGCAGAACTGAGAAATGCTTGGCCGAGGCGGCTGTGAGATACAGAGTATACC 821
QY 221 SerAlaValaAspLeuTrpSerLeuGlyIleIleLeuLeuGlnIlePheSerGlyMetLys 240
DB 822 TCGAGCTGTGATCTCTGAGGCTTGGAGATCATTTTACTGGAATTTCTTCAGAAAGAAA 881
QY 241 LeuLysHisThrValaArgSerGlnGluTYrLysAlaAsnSerSerAlaIleIleAspHis 260
DB 882 CTGAACATCATCAGATCTCAGAGGTGAGAAAGCAAGCTTCTGATATTTATTTGATCAT 941
QY 261 IlePheAlaSerLysAlaValaAlaAlaIleProAlaTYrHisLeuAlaGlnPheLeu 280

```

```

DB 942 ATATTGCCAGTAAAGCACTGTGATGCCGGAATTCACGCTATCAGACTTCAAGACCTT 1001
QY 281 IleLysSerMetLeuHisAspAspProSerArgAlaIleProAlaGluMetAlaLeuCyb 300
DB 1002 ATCAAAAGCATGCTTCATGACAGCCCAAGAGATCCCTGCTGAGATGGCTTGTGTC 1061
QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAlaPheValMetLeuProThr 320
DB 1062 AGCCCATCTTTAGATTCCTTTTTCCTCCCTCATATTTGAAGATGTGATGCTTCAACT 1121
QY 321 ProValLeuArgLeuLeuAlaValLeuAspAspArgTYrLeuGlyAlaGlnGluTYr 340
DB 1122 CAGTGCTCAGACTCTCTCATATGCTGATGATGATCTTCTTGAAGATGAGATGAATAT 1181
QY 341 GluAspValaValaGluAspValaLysGlnGlyCybaLysTYrGlyProValaSerLeu 360
DB 1182 GAAGATGTTGTGAAGATGTAAAGAGAGTGTCAAGATATGACCAAGTGGTTCCTCTG 1241
QY 361 LeuValProLysGluAsnProGlyArgGlnValaPheValGluTYrAlaAsnAlaGly 380
DB 1242 CTTGTTCAGAGAAATCTTGAGAGAGCAAGCTTTGTGAGTACGCAAAACCTGGA 1301
QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValaVal 400
DB 1302 GATTCGAAGCTGCTCAGAAATGCTGACTGGAGAGATCTTGAAGGAACTTTGTGTG 1361
QY 401 AlaThrPheTYrProLeuSerAlaTYrLysArgGlyTYrLeuTYrGlnThrLeuLeu 419
DB 1362 GGTACATTTACCCGCTGAGTGTCTTACAAAGAGGAGATATCTTATCAAACTTGCTT 1418

```

RESULT 11

ADQ24401

ID ADQ24401 standard; DNA; 3998 BP.

AC ADQ24401;

DT 26-AUG-2004 (first entry)

XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 7221.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;

XX de.

OS Homo sapiens.

PN WO2004048938-A2.

XX 10-JUN-2004.

PD 10-JUN-2004.

XX 26-NOV-2003; 2003WO-US038193.

PF 26-NOV-2002; 2002US-0429739P.

XX (PROT-) PROTEIN DESIGN LABS INC.

PA Aziz N, Ginsburg WM, Zlotnick A;

XX WPI; 2004-441208/41.

DR Early detection of soft tissue sarcoma comprises determining expression

XX of a gene in a first soft tissue sample and a normal soft tissue sample

PT and comparing the gene expression, also useful in treating soft tissue

PT sarcoma.

PS Example 2; SEQ ID NO 7221; 210bp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma

XX which comprises obtaining a first soft tissue sample from an individual

XX and a normal soft tissue sample from the same or different individual,

XX determining the expression of a gene in both samples and comparing the

XX expression of the gene in both soft tissue samples, where a higher level

XX of protein expression in the first soft tissue sample indicates the

XX presence of soft tissue sarcoma. The method of the invention has

CC cytoskeletal applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.

XX Sequence 3998 BP; 1066 A; 777 C; 885 G; 1255 T; 0 U; 15 Other;

Alignment Scores:

| Pred. No.: | 1.06e-190 | Length: | 3998 |
|------------------------|-----------|---------------|------|
| Score: | 2042.50 | Matches: | 395 |
| Percent Similarity: | 94.76% | Conservative: | 3 |
| Best Local Similarity: | 94.05% | Mismatches: | 21 |
| Query Match: | 92.59% | Indels: | 1 |
| DB: | 12 | Gaps: | 1 |

US-10-798-532-4 (1-419) x ADQ24401 (1-3998)

```
QY 1 MetAlAG1SerGIyCysAlATPG1yAlAG1uProArPhreLeuG1ua1aPheGIy 20
Db 160 ATGGCCGGATCCGGCTGCGCTGGGGCGCGAGCCCGCGTTTCTGAGGCTTCGGG 219
QY 21 ArgLeuTPG1uValG1nSerArG1eug1ySerGIySerSerAlaSerValTYrArGVal 40
Db 220 CGGCTGGGAGGTACAGAGCCGCTGGGTAGCGGCTCTCCGCTCGGCTGATCGGGTT 279
QY 41 ArgCyCysGIyVAsnProGIySerProProGIyAlaLeuArg1nPhleuP-roproGIy 60
Db 280 CGCTGTGGCGCAACCTGTGCTGCGCCCGCGGCCCTCAAGAGTCTTGGCGCGAGA 339
QY 61 ThrThG1yAla1a1a1aSerAla1a1aG1uTYrG1yPheArg1ySg1uArg1a1a1aLeu 80
Db 340 ACCACCGGGGCTCGGCTCTGCGCGAGATGTTTAAAGATGTTTCAACACCG 399
QY 81 GluG1nLeuG1nG1yHisArgAen1leValThrLeuTYrG1yValPheThr1leHis--- 99
Db 400 AAGGTCTGTGTGACCTTAAACGAGAGTGACTTTGTATGAGAGTGTTCACATCCAGGTT 459
QY 100 PheSerProAenValProSerArG1yVleuLeuLeuG1uLeuAenValSerValSer 119
Db 460 GTTCTCCAAAGTGTGCATCACCGCTGTGTTCTTGAACTCTTGATGTCAGTGTTCG 519
QY 120 GluLeuLeuLeuTYrSerSerHisG1nG1yCysSerMetTrpMet1leg1nHisCysAla 139
Db 520 GAATTCCTCTTATATTCAGTCCAGGGTGTTCATGATGATGATGATGATGATGATGATG 579
QY 140 ArgAspVal1eug1u1a1eua1aPheLeuHisG1uG1yTYrValHisAlaSerLeu 159
Db 580 CGAGATGTTTGGAGGCCCTTGTCTTTCATCATGAGGGCTATGTCATGCGAGCTC 639
QY 160 LysPProArgAen1leuLeuTPSerAlaG1uAenG1uCyPheLeuVal1eAenPheG1y 179
Db 640 AAACCGAGTAACTATGTGAGATGACAGAAATGAATGTTTAACTCATGTGACTTTGGA 699
QY 180 LeuSerPheLysG1uG1yAenG1uAenValLysTYr1leg1nThrAenP1yTYrArgAla 199
Db 700 CTTAGCTTCAAGAGAGGCAATCGAGATGTAAGTATATTCAGACAGCGGCTATCGGGCT 759
QY 200 ProG1uAlaG1uLeuG1nAenCySLeuAlaG1nAlaG1yLeuG1nSerArPThrG1uCyS 219
Db 760 CCAGAGACGAAATGCAAAATTCCTTGCCGAGGCTGCGCTGCAAGTATACAGATG 819
QY 220 ThrSerAlaValAspLeuTPSerLeuG1y1le1leuLeuG1uMetPheSerG1yMet 239
Db 820 ACCTCAGCTGTGATGTGTGAGACCTTAGAATCATTTTACTGAAATGTTTTCAGAGATG 879
QY 240 LysLeuLysHisThrValArgSerG1nG1uTPTrpLysAlaAenSerSerAla1le1leAenP 259
Db 880 AAACCTAAACATACAGTCAATTCAGAAATGAGAAACAGTTCCTGCTATATATGAT 939
QY 260 His1lePheAlaSerLysAlaValAlaSnAla1a1leProAlaTYrHisLeuArgAsp 279
```

```
Db 940 CACATATTTCCAGTAAAGCACTGGTGAATGCCGCAATTCAGCTTACCTAAGAGAC 999
QY 280 Leu1leLysSerMetLeuHisAspArProSerArG1y1leProAlaG1uMetAlaLeu 299
Db 1000 CTATATCAAAAGATCTTCATATATATCCAGAGAGATTCCTGCTGAATGGCATTG 1059
QY 300 CysSerProPhePheSer1leProPheAlaProHis1leg1uAenValMetLeuPro 319
Db 1060 TGCAGCCCATTTTATAGCATTCCTTTGGCCCTCATATTTGAAGATCTGGTATGCTCC 1119
QY 320 ThrProValLeuArgLeuLeuAenValLeuAenAspArTYrLeuG1yAenG1uG1u 339
Db 1120 ACTCAGAGCTTAAGACTGTGATGTGTGATGTATTAATCTTGAGAAATGAGAGGA 1179
QY 340 TYrG1uAenValValG1uAenValLysG1uG1uCySg1nLysTYrG1yProValAlaSer 359
Db 1180 TATGAAGATGTGTGAAGATTAAGAGAGGTCTCAAAATATATGACCACTGGTATCT 1239
QY 360 LeuLeuValProLysG1uAenProGIyArG1yG1nValPheValG1uTYrAlaSnAla 379
Db 1240 CTACTTGTCCAAAGAAATCTCGCAGAGGACAAGTCTTTGTTGAGATGCAATGCT 1299
QY 380 GlyAspSerLysAlaAlaG1nLysLeuLeuThrG1yArG1yMetPheAspG1yLysPheVal 399
Db 1300 GGTGATTCCAAAGCTCGCAGAAATTAAGTGAAGAGATGTTGATGAGGAACTTTGTT 1359
QY 400 ValAlaThrPheTYrProLeuSerAlaTYrLysArG1yTYrLeuTYrG1nThrLeuLeu 419
Db 1360 GTGGCTACATTTACCCGCTGAGTCTTCAAGAGGAGATATCTGATCAAACTTGCT 1419
RESULT 12
AAH70764
ID AAH70764 standard; cDNA; 529 BP.
AC AAH70764;
AT 19-SEP-2001 (fixet entry)
DT 19-SEP-2001 (fixet entry)
DX 19-SEP-2001 (fixet entry)
DE Human cervical cancer marker nucleic acid 2038.
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
OS Homo sapiens.
XX Homo sapiens.
XX WO200142467-A2.
XX 14-JUN-2001.
XX 08-DEC-2000; 2000WO-US033312.
XX 08-DEC-1999; 99US-0169681P.
XX 21-DEC-1999; 99US-0171350P.
XX 14-MAR-2000; 2000US-0189315P.
XX 12-MAY-2000; 2000US-0203791P.
XX 09-JUN-2000; 2000US-0210600P.
XX 21-JUL-2000; 2000US-0220114P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX WPI; 2001-375006/39.
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
XX for assessing and detecting compounds for treating the cancer.
XX Claim 1; Page 434; 1051p; English.
XX The invention relates to novel genes (AAH68727-AAH73303) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
XX polypeptides are useful: to assess if a patient is afflicted with
XX cervical cancer or has a pre-malignant condition; to monitor the
XX progression of cervical cancer or a premalignant condition in a patient;
```

CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy

XX Sequence 529 BP; 156 A; 102 C; 132 G; 139 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | Length: | 529 |
|------------------------|---------|-----------------|
| Score: | 789.00 | Matches: 159 |
| Percent Similarity: | 95.78% | Conservative: 0 |
| Best Local Similarity: | 95.78% | Mismatches: 5 |
| Query Match: | 35.77% | Indels: 2 |
| DB: | 4 | Gaps: 0 |

US-10-798-532-4 (1-419) x AAH70764 (1-529)

```

QY 220 ThSerAlaValAspLeuTPSerLeuGlyIleIleuLeuGluMetPheSerGlyMet 239
    |||
DB 26 ACCTCAGCTGTATCTGTGAGCCTTAGAATCAATTTACTGAAATGTTCTCAGAAATG 85
    |||
QY 240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
    |||
DB 86 AAACGAAACATACAGTCAGATCTCAGAGATGAGCAAAACGTTCTGCTATTATTGAT 145
    |||
QY 260 HisIlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTrpHisLeuArgAsp 279
    |||
DB 146 CACATATTTCCAGTAAAGCAGTGGTGAATGCCGCAATTCACGCTTATCCTTAAAGAGAC 205
    |||
QY 280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeu 299
    |||
DB 206 CTATACAAAGCATGCTCTCATGATGATCCAAAGCAAAATCTCTGCTGAAATGGCATTG 265
    |||
QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
    |||
DB 266 TGGAGCCCATCTTTTGGCATCTCTTTGCCCTCATATTAAGATCGTCATGCTTCC 325
    |||
QY 320 ThrProValLeuArgLeuLeuAsnValLeuAspAspTrpLeuGluAsnGluGlu 339
    |||
DB 326 ACTCCAGTCTAAGACTGCTGAATGTCTGATGATGATATTCCTTGAAGATGAAGAGAA 385
    |||
QY 340 TyrGluAsp-ValValGluAspValLysGlu-GluCysGlnLysTyrGlyProValValS 359
    |||
DB 386 TATGAGATTTGTTGAGATGTTAAAGAGAGGTCTCAAAATATGACCGGGGAT 445
    |||
QY 359 LeuLeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsn 379
    |||
DB 446 CTCTACTGTGTCAAAGAAATCTCGGCAAGAGAACAGTCTTGTGTGATGCAAAAG 505
    |||
QY 379 IacGlyAspSerLys 383
    |||
DB 506 CTGGGATTCMAAG 519
    |||

```

RESULT 13

AAA02535
 ID AAA02535 standard; cDNA; 722 BP.

AC AAA02535;

DT 19-MAY-2000 (first entry)

XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:2526.

XX Human; colon cancer; tumour; diagnosis; gene expression product; probe;

KW detection; cancerous state; metastasis; identification; breast cancer;

KW oestrogen receptor-positive breast cancer; therapy;

KW oestrogen receptor-negative breast cancer; lung cancer; ss.

OS Homo sapiens.

XX MO958675-A2.

XX 18-NOV-1999.

PF 13-MAY-1999; 99WO-US010602.
 XX 14-MAY-1999; 98US-0085426P.
 PR 15-MAY-1998; 98US-0085537P.
 PR 15-MAY-1998; 98US-0085696P.
 PR 21-OCT-1998; 98US-0105234P.
 PR 27-OCT-1998; 98US-0105877P.

XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.

PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J,
 PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kaasam A,
 PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I,
 PI Leshkowitz D, Kila D, Garcia V, Jones LW, Stache-Grain B,

XX WPI; 2000-126369/11.

XX Polynucleotide library used to determine cancerous states of mammalian
 PT cells.

XX Claim 1; Page 1020; 1097P; English.

CC AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
 CC libraries constructed from human colon cancer cell lines. The present
 CC invention also describes a method of detecting differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell, comprising
 CC detecting at least one differentially expressed gene product in a test
 CC sample derived from a cell suspected of being cancerous, where detection
 CC of the differentially expressed gene product is correlated with a
 CC cancerous state of the cell from which the test sample was derived. The
 CC polynucleotide sequences can be used in a method for detecting
 CC differentially expressed genes correlated with a cancerous state of a
 CC mammalian cell. The polynucleotides can also be used as probes for
 CC detecting and mapping related genes. They can be used in diagnosis and
 CC prognosis of diseases and disorders (e.g. identification of pre-
 CC metastatic or metastatic cancerous states, stages of cancer, or
 CC responsiveness of cancer to therapy). This is particularly for breast
 CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
 CC negative breast cancer, lung cancer, and colon cancer

XX SQ Sequence 722 BP; 215 A; 129 C; 164 G; 193 T; 0 U; 21 Other;

Alignment Scores:

| Pred. No.: | Length: | 722 |
|------------------------|---------|-----------------|
| Score: | 766.50 | Matches: 182 |
| Percent Similarity: | 84.86% | Conservative: 3 |
| Best Local Similarity: | 83.49% | Mismatches: 20 |
| Query Match: | 34.75% | Indels: 14 |
| DB: | 3 | Gaps: 1 |

US-10-798-532-4 (1-419) x AAA02535 (1-722)

```

QY 152 GUGUUYTTRValHAlaAspLeuLysProArgAsnIleLeuTrpSerAlaGluAsnGlu 171
    |||
DB 1 GAGGGCTATGTCATGCGGNCCTCAAAACNACCTAACATATGTGAGTGCAGAGAAATGA 60
    |||
QY 172 CysPheLysLeuIleAspPheGlyLeuSerPheLysGluGluAsnGluAspValLysTyr 191
    |||
DB 61 TGTTTTAAATCATTTGACTTGTGACTTANCTTCAAAAGAGCAATCAGATGTAAGATAT 120
    |||
QY 192 TGGTTHrAspGlyTyrArgAlaProGluAlaGluLeuGlnAsnCysLeuAlaGlnAla 211
    |||
DB 121 ATTTCAGACGACGGGTATCGGCTCCAGAAACG-GAATTGCAAAATGCTTGGCCCANCT 179
    |||
QY 212 GlyLeuGlnSerAspDhGluCysThrSerAlaValAlaAspLeuTrpSerLeuGlyIleIle 231
    |||
DB 180 GGCTTCGACAGATGATCAAGATGATCTCAGCTGTGTGATCTGTGAGAGCTTGAATCATTT 239
    |||
QY 232 LeuLeuGluMetPheSerGlyMetLysLeuLysHisThrValArgSerGlnGluTrpLys 251
    |||
DB 240 TTAATCGAAATTTCTCAGGAATGAATCACTGAACATATCAGTCAATTCAGGAATGAAG 299
    |||

```

QY 252 AlaAsnSer-SerAlaIleIle-AspHisIlePheAlaSerIysAla--ValValAsnAl 270
 DB 300 GCAACAGTTTCGCTATTATTGATCACTATTTCACGTAAGCAANTGCTGATGC 359
 QY 270 aAlaIleProAlaTyHisIleuArgAspLeuIleIysSerMetLeuHisAspAsp-ProG 290
 DB 360 CCGAATTCACGCTATCACCATTAAAGACCTTATCAAAAGACATGCTTCATGATATCCAA 419
 QY 290 eArg-ArgIleProAlaIleMetAlaIleu-CysSerProPhePheSer-IleProPheA 309
 DB 420 GCGAGGAGAAATTTCTNCTGAAATGCGATTGGTCANCCATTCTTACGCAATTCCTTTTG 479
 QY 309 lAProHisIleIleGluAspLeuValMetLeuProThrProVal-LeuArgLeuLeuAsnVal 328
 DB 480 CCCCTCATATTGAAGATCTGNTCATGCTTCCACTCAGTGGCTAAGACTGCTGAATGTG 539
 QY 329 Leu-AspAspAspPyrLeuGluIysAnglu--GluGluTyTrGluAspValValGluAspVal 347
 DB 540 CTGGGNTGATGATATCTTGAGAAATGAAAGACGATTATGATGTTGTGAAGATGNT 599
 QY 348 LysGluGluCysGlnLys---TyRGLyPro 356
 DB 600 AAAAGAAAGATGCGCAAAATTTTGACCA 629

RESULT 14
 AAH71966
 ID AAH71966 standard; cDNA; 470 BP.
 AC AAH71966;

19-SEP-2001 (first entry)

Human cervical cancer marker nucleic acid 3240.

Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

Homo sapiens.

MO200142467-A2.

14-JUN-2001.

08-DEC-2000; 2000WO-US033312.

08-DEC-1999; 99US-0169681P.

21-DEC-1999; 99US-0171350P.

14-MAR-2000; 2000US-0189315P.

12-MAY-2000; 2000US-0203791P.

09-JUN-2000; 2000US-0210600P.

21-JUL-2000; 2000US-0220114P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Deeds J, Berger A, Zhao X;

WPI; 2001-375006/39.

New isolated nucleic acid for diagnosing and treating cervical cancer and for assessing and detecting compounds for treating the cancer.

Claim 1; Page 636; 1051pp; English.

The invention relates to novel genes (AAH68727-AAH73383) associated with cervical cancer with cytostatic activity. The nucleic acids and encoded polypeptides are useful: to assess if a patient is afflicted with cervical cancer or has a pre-malignant condition; to monitor the progression of cervical cancer or a premalignant condition in a patient; and to select and/or assess the efficacy of a compound or therapy for inhibiting cervical cancer in a patient. The nucleic acids may also be useful for gene therapy

Sequence 470 BP; 140 A; 91 C; 114 G; 125 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,966-62 Length: 470
 Score: 729.00 Matches: 147
 Percent Similarity: 96.71% Conservative: 0
 Best Local Similarity: 96.71% Mismatches: 3
 Query Match: 33.05% Indels: 2
 DB: 4 Gaps: 0

US-10-798-532-4 (1-419) x AAH71966 (1-470)

QY 220 ThrSerAlaValAspLeuTrpSerLeuGlyIleIleLeuGluMetPheSerGlyMet 239
 DB 18 ACCTCAGCTGTTGATCTGGAGCTTAGAAATCTTATCTGGAATGTCTCAGAAATG 77
 QY 240 LysLeuLysHisIleThrValArgSerGlnIuTrpLysAlaAsnSerSerAlaIleIleAsp 259
 DB 78 AAACGAAACATACAGTCATGATCTCAGAGATGAAAGGCAACAGATTCGTCTATTATGAT 137
 QY 260 HisIlePheAlaSerIysAlaValValAsnAlaAlaIleProAlaTyHisIleuArgAsp 279
 DB 138 CACATATTTCGCGATTAAGCAGTGTGAATGCCGCAATTCAGCCTATCACCTTAAGAGAC 197
 QY 280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaIleu 299
 DB 198 CTTATCAAAAGCATGCTTCATGATGATCCAGCAAGAAATTCCTGGAATGCGCAATTG 257
 QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
 DB 258 TGCAGCCCATTTCTTACGATTCCTTTGGCCCTCATATTAAGATCTGCTATGCTTCCC 317
 QY 320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspPyrLeuGluIysAngluGlu 339
 DB 318 ACTCCAGTCTAAAGCTGCTGAATGTGCTGATGATGATTAATCTTGAGATGAAGAGGA 377
 QY 340 TyTrGluAsp-ValValGluAspValIysGlu-GluCysGlnLysTyTrGluProValValS 359
 DB 378 TATGAAGATTTGTTGTGAAGATGTAAAGAGGAGCTGTCAAAATATATGACACGAGGAT 437
 QY 359 eLeuLeuValProLysGluAsnProGlyArg 369
 DB 438 CTCTACTGTGTCAAAGAAATTCCTGCGACA 469

RESULT 15
 AAH69101

ID AAH69101 standard; cDNA; 461 BP.

AC AAH69101;

19-SEP-2001 (first entry)

Human cervical cancer marker nucleic acid 375.

Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

Homo sapiens.

MO200142467-A2.

14-JUN-2001.

08-DEC-2000; 2000WO-US033312.

08-DEC-1999; 99US-0169681P.

21-DEC-1999; 99US-0171350P.

14-MAR-2000; 2000US-0189315P.

12-MAY-2000; 2000US-0203791P.

09-JUN-2000; 2000US-0210600P.

21-JUL-2000; 2000US-0220114P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Deeds J, Berger A, Zhao X;

DR WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 168-169; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

XX Sequence 461 BP; 134 A; 91 C; 109 G; 127 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 2,87e-53 | Length: | 461 |
|------------------------|----------|---------------|-----|
| Score: | 636.00 | Matches: | 138 |
| Percent Similarity: | 93.33% | Conservative: | 2 |
| Best Local Similarity: | 92.00% | Mismatches: | 5 |
| Query Match: | 28.83% | Indels: | 6 |
| DB: | 4 | Gaps: | 0 |

US-10-798-532-4 (1-419) x AAH69101 (1-461)

```
QY 220 ThSerAlaValAspLeuTpsSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMet 239
    |||||
DB 22 ACCTCAGCTGTTATCTGTGAGACCTTAGAATCATTTTACTGAAATGTTCTCAGGAATG 81
    |||||
QY 240 LysLeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAlaIleIleAsp 259
    |||||
DB 82 AAACCTGAACATACAGTCAGATCTCAGAGATGAGCAAAACAGTTCTGCTATTATTGAT 141
    |||||
QY 260 HisIlePheAlaSerLysAlaValAlaAsnAlaIleProAlaTyrHisLeuArgAsp 279
    |||||
DB 142 CACATATTGCTCCAGTAAGCAGTGTGAATGCCGCAATTCAGCTATCCCTAAGAGAC 201
    |||||
QY 280 LeuIleLysSerMetLeuHisAspAspProSerArgArgIleProIleGluMetAlaLeu 299
    |||||
DB 202 CTTATCAAAAGCATGCTTCTCATGATGCCAAGCAGAAATTCCTGCTGAATGSCATTG 261
    |||||
QY 300 CysSerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuPro 319
    |||||
DB 262 TGCAGCCCAATCTTTAGCATTCCTTTGCCCTCATAT-GAAGATCTGTCATGCTTCCC 320
    |||||
QY 320 ThrProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGluAsnGluGlu 339
    |||||
DB 321 ACTCCAGTGTAGACTGCTGAATGCTGATGATGATGATGATGATGATGATGATGATGATG 380
    |||||
QY 340 TyrGluAsp--ValValGluAspValLysGlu--GluCysGlnLysTyrGlyPro-ValY 358
    |||||
DB 381 TATGAAGATTGTGTGAAGATGTAAAGAGGAGAGGTGTCAAAAATATGACACAGTGG 440
    |||||
QY 358 alSerLeuLeuValProLys 364
    |||||
DB 441 TATCTCTACTTGTGCCAA 460
```

Search completed: November 29, 2004, 11:42:36
Job time : 520 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 26 Seconds

(without alignments)
1550.570 Million cell updates/sec

Title: US-10-798-532-4

Perfect score: 2206

Sequence: 1 MAGSGCAGMGPFRFLFAGF.....VATFYPLSAVKRYLYQTLL 419

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------------|
| 1 | 1052 | 47.7 | 200 | 2 | 148615 gene KIS protein - |
| 2 | 248.5 | 11.3 | 948 | 2 | T24445 hypothetical prote |
| 3 | 240.5 | 10.9 | 720 | 2 | A96807 hypothetical prote |
| 4 | 240.5 | 10.9 | 826 | 2 | JC5153 mitogen-activated |
| 5 | 240 | 10.9 | 388 | 2 | S48879 protein kinase SKR |
| 6 | 238.5 | 10.8 | 387 | 2 | T37758 protein kinase skp |
| 7 | 234.5 | 10.6 | 433 | 2 | S37790 probable serine/th |
| 8 | 233 | 10.6 | 369 | 2 | A56492 protein kinase ERK |
| 9 | 232.5 | 10.5 | 367 | 2 | S68680 stress-activated p |
| 10 | 231.5 | 10.5 | 354 | 2 | T45138 protein kinase HOG |
| 11 | 231 | 10.5 | 435 | 2 | S64950 protein kinase HOG |
| 12 | 230 | 10.4 | 298 | 1 | A44878 protein kinase (EC |
| 13 | 230 | 10.4 | 298 | 2 | A41327 protein kinase (EC |
| 14 | 230 | 10.4 | 326 | 2 | S23587 protein kinase (EC |
| 15 | 230 | 10.4 | 373 | 2 | T13024 probable protein k |
| 16 | 230 | 10.4 | 376 | 2 | S40470 mitogen-activated |
| 17 | 230 | 10.4 | 376 | 2 | S40471 mitogen-activated |
| 18 | 230 | 10.4 | 427 | 2 | JC5693 stress-activated p |
| 19 | 228 | 10.3 | 423 | 2 | S43968 p54-alpha-2 stress |
| 20 | 227 | 10.3 | 423 | 2 | S43967 p54-alpha stress-a |
| 21 | 227 | 10.3 | 424 | 2 | S71102 protein kinase JNK |
| 22 | 227 | 10.3 | 529 | 1 | S71774 calcium-dependent |
| 23 | 226 | 10.2 | 406 | 1 | KIHUCT phosphorlyase kina |
| 24 | 225.5 | 10.2 | 301 | 1 | SI19209 protein kinase (EC |
| 25 | 225.5 | 10.2 | 416 | 2 | A48249 pre-mRNA splicing |
| 26 | 225 | 10.2 | 371 | 2 | S60121 mitogen-activated |
| 27 | 224.5 | 10.2 | 335 | 2 | T23050 hypothetical prote |
| 28 | 224.5 | 10.2 | 342 | 2 | T21098 protein kinase (EC |
| 29 | 224.5 | 10.2 | 371 | 2 | T14915 mitogen-activated |

| | | | | | |
|----|-------|------|-----|---|---------------------------|
| 30 | 224.5 | 10.2 | 469 | 1 | S17656 Ca2+/calmodulin-de |
| 31 | 223.5 | 10.1 | 367 | 2 | JC5252 mitogen-activated |
| 32 | 223.5 | 10.1 | 426 | 2 | S43969 p54-beta stress-ac |
| 33 | 223.5 | 10.1 | 464 | 2 | S71104 protein kinase JNK |
| 34 | 223.5 | 10.1 | 678 | 2 | T43539 spindle checkpoint |
| 35 | 223 | 10.1 | 346 | 1 | T48157 protein kinase (EC |
| 36 | 223 | 10.1 | 406 | 2 | D84898 probable mitogen-a |
| 37 | 223 | 10.1 | 427 | 2 | JC5694 stress-activated p |
| 38 | 223 | 10.1 | 575 | 2 | JC7794 lammer kinase homo |
| 39 | 223 | 10.1 | 690 | 2 | T38052 probable protein k |
| 40 | 222.5 | 10.1 | 301 | 1 | A48041 protein kinase (EC |
| 41 | 222.5 | 10.1 | 424 | 2 | A55480 c-Jun amino-termin |
| 42 | 221 | 10.0 | 358 | 2 | S23383 protein kinase (EC |
| 43 | 220.5 | 10.0 | 382 | 2 | JC5531 c-Jun amino-termin |
| 44 | 220.5 | 10.0 | 602 | 2 | S60052 calcium-dependent |
| 45 | 220 | 10.0 | 371 | 2 | T09622 protein kinase RMK |

ALIGNMENTS

```

RESULT 1
148615
Gene KIS protein - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: 148615
R/Maucuer, A.; Camonis, J.H.; Sobel, A.
Proc. Natl. Acad. Sci. U.S.A. 92, 3100-3104, 1995
A/Title: Stathmin interaction with a putative kinase and coiled-coil-forming protein dome
A/Reference number: 148282; PMID:95241452; PMID:7724523
A/Accession: 148615
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-200 <RES>
A/Cross-references: UNIPROT:P97343; EMBL:X82320; NID:G791078; PIDN:CAA57763.1; PID:G791078
A/Genetic:
A/Genes: KIS

Query Match      47.7%; Score 1052; DB 2; Length 200;
Best Local Similarity 99.0%; Pred. No. 8.1e-54;
Matches 198; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 122 LLYSSHQSGSMWMIQHCAADVLEALAFLEHGGVYHADLPKRNILMSAENECFKLIDFGLS 181
DB 1 LLYSSHQSGSMWMIQHCAADVLEALAFLEHGGVYHADLPKRNILMSAENECFKLIDFGLS 60

QY 182 FRKGNQDVXYIQTQGYRABEALONCLAQAGLSPTCTSAVDLWSLGIILEMFGMKL 241
DB 61 FRKGNQDVXYIQTQGYRABEALONCLAQAGLSPTCTSAVDLWSLGIILEMFGMKL 120

QY 242 KRTVNSQEWKANSALIIDIFASKAVVNAIIPAYHLRDLIKSMHDDPERRIPAEWALCS 301
DB 121 KRTVNSQEWKANSALIIDIFASKAVVNAIIPAYHLRDLIKSMHDDPERRIPAEWALCS 180

QY 302 PPFSTPFAPIHIDVWLPTR 321
DB 181 PPFSTPFAPIHIDVWLPTR 200

RESULT 2
T24445
hypothetical protein T04C10.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T24445
R/Burton, J.
submitted to the EMBL Data Library, March 1996
A/Reference number: Z19891
A/Accession: T24445
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-948 <WIL>

```

A:Cross-references: UNIPROT:Q9MOL7; EMBL:Z69885; PIDN:CAA93756.1; GSPDB:GN00028; CESP:TD
 A:Experimental source: clone T04C10
 C:Genetic:
 A:Gene: CESP:T04C10.1
 A:Map position: X
 A:Intons: 64/3; 181/3; 214/3; 238/1; 272/3; 331/2; 391/2; 457/1; 726/3; 830/2; 923/3

Query Match 11.3%; Score 248.5; DB 2; Length 948;
 Best Local Similarity 25.7%; Pred. No. 5.6e-07;
 Matches 104; Conservative 57; Mismatches 113; Indels 131; Gaps 17;

QY 88 NYTLVGVTFTHSPVPSRCLLELDVSVSELLYSHGCGSMWIGHQCARDVLEALA 147
 DB 455 NYTLVGVTFTHSPVPSRCLLELDVSVSELLYSHGCGSMWIGHQCARDVLEALA 147
 QY 148 PLHNE--GYVHADLIDFGLSPFEGNODVYIOTD--GYRADEAEI 204
 DB 510 FLSSPELSIITHCDLKENVLLVNAKRSQIRVIDFGSSCQIGRIYQIQSRFRSEVLL 569
 QY 205 QNCLAQAGLSQDTECTSAVDLWSLGIILLEMFGS-----MKL----- 241
 DB 570 -----GLAYDVK-----IDMWSLGCILVEMHGEPLFAGSSEVDQMKLVEVLGMPKX 618
 QY 242 -----KHTYRSQEWKANSATIDHIFASCAVYNAAI 272
 DB 619 MLDIGPKTHKYPDKTEGDIYVCKKTRDGYRHT-----YKAPGARKLHEILG-----VTSQG 669
 QY 273 PA-----YHLRDLKSMLEHDPSSRRIPAEMLGSPFSI--PFAPIHEDLV 316
 DB 670 PGRRLIGEGHSEVDYSKFKDLKRMLODPDKRI-----SPYVVRHPLFKQKEEV 722
 QY 317 MLPTPVLRLNLVLDLDYLGNEEYEDVEDVKECCQY--GPVSLVLPKENGRCQVTV 374
 DB 723 PSQPEVSH-----SNLQOQQLYIQPSSQMSQVWESVSGSYV 762
 QY 375 E----YANAGDSQAQKLTLGRMPDGF--VATFPPLSLYKRGY 413
 DB 763 EDNGMYROAPGSSANPISVTSPDEGDAMEVDAGRRRRFSAHQNY 807

RESULT 3

A:Cross-references: UNIPROT:Q9MOL7; EMBL:Z69885; PIDN:CAA93756.1; GSPDB:GN00028; CESP:TD
 A:Experimental source: clone T04C10
 C:Genetic:
 A:Gene: CESP:T04C10.1
 A:Map position: X
 A:Intons: 64/3; 181/3; 214/3; 238/1; 272/3; 331/2; 391/2; 457/1; 726/3; 830/2; 923/3
 Query Match 11.3%; Score 248.5; DB 2; Length 948;
 Best Local Similarity 25.7%; Pred. No. 5.6e-07;
 Matches 104; Conservative 57; Mismatches 113; Indels 131; Gaps 17;
 A:Cross-references: UNIPROT:Q9CA22; GB:AE005173; NID:96437532; PIDN:AAF08564.1; GSPDB:GN
 A:Genetic:
 A:Gene: T32E8.5
 A:Map position: 1

Query Match 10.9%; Score 240.5; DB 2; Length 720;
 Best Local Similarity 25.7%; Pred. No. 1.2e-06;
 Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;
 QY 11 EPPRFLEAFGRIMOVOSRSGSSASAYVRRCGNGSPGALROPLPPTTTAAASAAE 70
 DB 331 DPLDFKVGKGLYQRLGKIGSGSSSEVHKV---SSDCTTYALKKIKLGRDYATA---- 383

QY 71 YGFRKERAALEQLOGRNIVTLVGVTFTHSPVPSRCLLELDVSVSE----- 120
 DB 384 YGFCOEIGYLLKTKKKTNIQL-----IDYEVDKTLLEVLNGTSMNKGKRVKEDGF 436

QY 121 ---LLLYS-----SHGCGSMW-----MIGHCARDVLEALHHEGYHADLK 160
 DB 437 IYVLEGEIDLAHNLQKWRKEIEGSDRTIDENWLRFPYQQLQAVNTTHERRYHSDLK 496
 QY 161 PNITWASNEBCFKLIDFGLSPFEGNODVYIOTD-----GYRADEALQNCCLAQGLQ 214
 DB 497 PANFL--LVRGFLKILIDGFLA--KAINSDPTNIQRSSQGLTSSYMPAFMCHESDENG-- 551
 QY 215 SDTECTSAVDLWSLGIILLEMFGSMKLKHTYRSQEWKANSATIID--HIFASKAVYNAAI 272
 DB 552 NTIKCGRPEDISLGCILYQWYVG--RTPADYKTFW--AKFKVITDPNHETYNQLSNP-- 607
 QY 273 PAVHLRDLKSMLEHDPSSRRIPAEMLGSPFSI--PFAPIHEDLV 316
 DB 608 ---WLDLWKKCLAWDRQWRIPPELLQHPFLAPPI--PHEPOVKTIKLSLIAESGSD 663
 QY 312 -----IBDLVMLPTPVLRLNLVLDLDYLGNEEYEDVEDVKECC 351
 DB 664 DKAWSISGLLEGLLSNPAPLPR-----NDVLDSRDQQLLSRVSLELC 706

RESULT 4

A:Cross-references: UNIPROT:Q9MOL7; EMBL:Z69885; PIDN:CAA93756.1; GSPDB:GN00028; CESP:TD
 A:Experimental source: clone T04C10
 C:Genetic:
 A:Gene: CESP:T04C10.1
 A:Map position: X
 A:Intons: 64/3; 181/3; 214/3; 238/1; 272/3; 331/2; 391/2; 457/1; 726/3; 830/2; 923/3
 Query Match 10.9%; Score 240.5; DB 2; Length 826;
 Best Local Similarity 25.1%; Pred. No. 1.4e-06;
 Matches 103; Conservative 57; Mismatches 145; Indels 105; Gaps 19;
 A:Cross-references: UNIPROT:Q94656; EMBL:X82646
 C:Keyword: phosphotransferase

QY 23 KQVOSRLGSSASAYVRRCGNGSP-----GALRQFLPPTCTGAASAAEYGRKER 77
 DB 23 YDLKKVKGAGVGVFKGCKKKNIVAVKIFGAFON-----CTDAQRTFR--EI 71
 QY 78 AALBOLQGRNIVTLVGVTFTHSPVPSRCLLELDVSVSEL-----LYSSHGCGSMW 133
 DB 72 IFLVELNGHDNIKLMGVK--AKNDNDIYILPFMETDHEVIAKADLEIEHK----- 123
 QY 134 MIGHCARDVLEALHHEGYHADLKPRNITWASNEBC--FKLIDFGLS-----FKSGN 186
 DB 124 --KYIITVQLRAKTIHSGGLIHRDIKPSNII--VNSECHKIVADFGIARSISTVHNEK 179
 QY 187 QDV--KYIOTDGYRAPEALQNCCLAQGLQSTECTSAVDLWSLGIILLEMFGSMKL--- 241
 DB 180 VDLIDVYATVTRVRAPEI-----LGSHTYEDVDWMSLSCINGELCGKPLFTG 229
 QY 242 KHTVRSQ-----WKANSATIID--HIFASKAVYNA--AIPAVYHR-----DLI 281
 DB 230 NSTWQLEKIITVIGKPKKOIEDIRSPAEKTISSFPVLDKKNLKDICYKASNESLDLL 289
 QY 282 KSMLEHDPSSRRIPAEML-----CSPFSIPFAPIHEDLVMLPTPVLRLNLVLDLDYLGNE 337
 DB 290 EKLQGFNPSKRISAEALGKRYVEEFHSIIDEPTCRHIIITP----- 331
 QY 338 EYEDVVEVKECCQYRGVPSLLVPKENRGQYFVEYANAGDSKAQK 387

Db 332 -----INDNTKYRVNFRVNVVFEVIMRNKFNHNVL-----NOGESKKEK 372

RESULT 5

S48879
protein kinase SMK1 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YP9499.10; protein YPR054w
C/Species: Saccharomyces cerevisiae
C/Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C/Accession: S48879; S54076
R/Kritak, L.; Strich, R.; Winters, R.S.; Hall, J.P.; Mallory, M.J.; Kreitzer, D.; Tuan, Genes Dev. 8, 2151-2161, 1994
A/Title: SMK1, a developmentally regulated MAP kinase, is required for spore wall assembly
A/Reference number: S48879; PMID:95047362; PMID:7958885
A/Accession: S48879
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-388 <RRI>
A/Cross-references: UNIPROT:P41808; EMBL:L35047; NID:G530516; PIDN:AB59325.1; PID:G5385
R/Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A/Reference number: S54059
A/Accession: S54076
A/Residues: 1-388 <BAD>
A/Molecule type: DNA
A/Cross-references: EMBL:249219; NID:G805025; PIDN:CAA89172.1; PID:G805035; MIPS:YPR054w
A/Experimental source: strain AB972
C/Genetics:
A/Genes: SGD:SMK1
A/Cross-references: SGD:S0006258; MIPS:YPR054w
A/Map position: 16R
C/Superfamily: kinase-related transforming protein; protein kinase homology
C/Keywords: ATP; phosphotransferase; signal transduction
F/36-306/Domain: protein kinase homology <KIN>
F/44-52/Region: protein kinase ATP-binding motif

Query Match 10.9%; Score 240; DB 2; Length 388;

Best Local Similarity 25.8%; Pred. No. 6,6e-07;
Matches 103; Conservative 61; Mismatches 133; Indels 102; Gaps 19;

QY 21 RLWQVSRIGSGSSAVYRVRCGNGPSPGALRQFLPPTGAAASAAEYG----- 72
Db 10 RAINVNASNIGAPQQRITFAKERISIPGY--YRIIQFGKAGVGVCSVFKGRSPARLA 67
QY 73 -----FKKE---RAALBQL-----QGHRTVTLVGVTFTHFSNVPSCILLLEL 116
Db 68 VKKISNIFENKEIILKRAIRBLKFMNPFKGNIVNLIDETIVSSPDGLYC-YQSLIDY 126
QY 117 SVSELLLYSSHQCSMMIIOHCARDVLEALAFHHEGYVHADLKPRNIIWLSAENECFKLI 176
Db 127 DLAK-VIHSSVQ-LSEFHNIKVFYQILGLKXIHSDVIRDLKPGNIICTL-NGCLKIC 183
QY 177 DFGLS-----FKENQDVK-----YIQDGYRPAEAEIIONCIAQAGLOSDECTSAVDL 225
Db 184 DFLANGIHAGFKCKSTVQPHITNVVATRWYRABEL-----LISNQYSSKSVDI 233
QY 226 WSGIITILE-----WESGKILKHTV-----RSQEW---KANSAAI 257
Db 234 WAGCILAIEFYARKPVFMGDSMHQIFELIKVLGTPDKDILIKFTIKANLAKSNINPV 293
QY 258 IDH1-----FASKVAVNAAIPIAYHLRLDIKSMLEHDPERRIPAEWALCSPPFSIPFAP 310
Db 294 YKTIPIINSNIFPASHAIN-----LIESLIMHDSIHRILNVEQALSHPLINVRKP 343
QY 311 HIEDLVMLPTPVRLNLNVLDDYLGNEEYED-VVEDVK 348
Db 344 D-DEPVCLQGP-----FDFTYESELNMSKLDYLVVEVK 377

RESULT 6
T37758
protein kinase skip1 - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: T37758

R/Wedler, H.; Wambutt, R.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, December 1998

A/Reference number: 221744

A/Accession: T37758

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-387 <MED>

A/Cross-references: UNIPROT:Q10452; EMBL:AL035064; PIDN:CAA22609.1; GSPDB:GN00066; SPDB:4

A/Experimental source: strain 972h-1; cosmid c1687

C/Genetics:
A/Genes: SPDB:SPAC1687.15

A/Map position: 1

A/Introns: 7/1; 70/3

C/Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 10.8%; Score 238.5; DB 2; Length 387;

Best Local Similarity 24.6%; Pred. No. 8,1e-07;

Matches 78; Conservative 48; Mismatches 112; Indels 79; Gaps 10;

QY 29 LGGSSASVYRVRCGNGPSPGALRQFLPPTGAAASAAEYGRKRAALEQLQGHRT 88
Db 38 VSGSGFVWQVHLIBDSK--AAIKRVL-----QDKFRKRELQIMRKHPN 84
QY 89 IYTLGVF-----TIHSPNV---PSRCILLLELDYVSVELLYSSHQGS 131
Db 85 IYDLIAYYTTGDNDEVIYANLVLEFMPETIRASRLYTRQLSPMIEVLKY----- 137
QY 132 MMHIOHCARDVLEALAFHHEGYVHADLKPRNIIWLSAENECFKLIDFGLS--FKENQDV 189
Db 138 -----IYQLRSIAYIHASGICHRDIPQMLDPENGILKLDGFSAKILVAGEPNV 190
QY 190 KYIQDGYRPAEAEIIONCIAQAGLOSDECTSAVDLMSGIITILEMFGS----- 238
Db 191 SYICSRYYRABEL-----IFGADYTHAIDIWSTGCVMAELMGHPLPFGESGID 240
QY 239 -----MKLHGYRSQSEWKNSSAIIIDHIFAS-----KAVVNAAIPIAYHLRLDIKSMLEHD 287
Db 241 QLVEIILKILGTSRQKIKTMNPNYWEHRRPQIRPQPLSRVFSRSPDLAL-DLSKMLQY 299
QY 288 DPSRRIPIAEWALCSPPF 304
Db 300 TPTDRLTAABAMCHPFF 316

RESULT 7

S37790
probable serine/threonine-specific protein kinase (EC 2.7.1.-) YKL161c - yeast (Saccharom

N/Alternate names: protein YKL615

C/Species: Saccharomyces cerevisiae

C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Aug-2004

C/Accession: S37790; S37991; S44567

R/Vandenbol, M.; Bolle, P.; Dion, C.; Portetelle, D.; Hilger, F.

submitted to the EMBL Data Library, September 1993

A/Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loc

A/Reference number: S37786

A/Accession: S37790

A/Molecule type: DNA

A/Residues: 1-433 <VAN>

A/Cross-references: UNIPROT:P36005; EMBL:Z26877; NID:G407482; PIDN:CAA81493.1; PID:G40746

A/Experimental source: strain S288C

R/Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.

submitted to the Protein Sequence Database, March 1994

A/Reference number: S37976

A/Accession: S37991

A/Molecule type: DNA

A/Residues: 1-433 <VA2>

A/Cross-references: EMBL:Z28161; NID:G486280; PIDN:CAA82003.1; PID:G486281; MIPS:YKL161C

A/Experimental source: strain S288C

R/Vandenbol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.

Yeast 10, 35-40, 1994

A/Title: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAP4 loci of c

Qy 365 E 365
Db 356 Q 356

RESULT 10

T45138
protein kinase ekpl [imported] - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C/Accession: T45138
R/Player, S.E.; Feoktistova, A.; Burke, J.D.; Woodgett, J.R.; Gould, K.L.
Mol. Cell. Biol. 16, 179-191, 1996
A/Title: Schizosaccharomyces pombe ekpl+ encodes a protein kinase related to mammalian g
A/Reference number: 2252927; MIMD:96104567; PMID:8524294
A/Accession: T45138
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-354 <PLY>
A/Cross-references: EMBL:L29449; NID:g1914882; PIDN:AAB51081.1; PID:g1914883
C/Genetics:
A/Gene: ekpl
A/Intons: 6/3
C/Function:
A/Pathway: cells w
A/Note: cells with deletions in ekpl are sensitive to heat shock and exhibit defects in
C/Keywords: phosphoprotein, phosphotransferase
F/335/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 10.5%; Score 231.5; DB 2; Length 354;
Best Local Similarity 24.3%; Pred. No. 1.8e-06;
Matches 77; Conservative 48; Mismatches 113; Indels 79; Gaps 10;

Qy 29 LSGSSASVYVRCCGPPGALROFLPPGTGAASAAYGFKERAALEFOLQGHN 88
Db 38 VGSFGVVMQVVALISDSK-AAIKRVL-----QDKRFKRELQIRIMNDPV 84
Qy 89 IYLVYGVF-----TIHSPNV---PSRCLLELLDVSELLSYSHQGS 131
Db 85 IYDLVYVYTTGDNSEVYINLVLEPPEITYPASRLYTRQKLSMPLEVKY----- 137
Qy 132 MMVIOHCARDVLEALAFHHEGYVHADLKPRNIIWSENCFKIDFGLS--FKEGNQDV 189
Db 138 -----IYQLASLAVIHASGICHRDIPQNLILDPENKILKCPFGSAKILVABPNV 190
Qy 190 KYIQDGYRAPAEALQNCIAQAGLQSDTECTSAVDLWSLIGIILLEMFSG----- 238
Db 191 SYICSRVYRAPEL-----IFGATDYTHAIDIMSTGCVABMLGHPLPFGESGD 240
Qy 239 -----MKLKITVASQEMKANSALIDHIFAS-----KAVNNAIPAYHLRDLIKSMHD 287
Db 241 QLVETIKILGTPSREQIKTNPNYMERFQIRPPLSRVSRSPVLDAI-DLISKMLQY 299
Qy 288 DPSRIIPAEWALCPFF 304
Db 300 TPTDRILTAEMAKCHPFF 316

RESULT 11

S64950
protein kinase HOG1 (EC 2.7.1.-), mitogen-activated - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein l2931; protein YLR113w
C/Species: Saccharomyces cerevisiae
C/Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 16-Aug-2004
C/Accession: S64950; S42145; S69400
R/Verhaaselt, P.; Voet, M.; Volckaert, G.
submitted to the Protein Sequence Database, May 1996
A/Reference number: S64443
A/Accession: S64950
A/Molecule type: DNA
A/Residues: 1-435 <VER>
A/Cross-references: UNIPROT:P32485; EMBL:Z73285; NID:g1360507; PIDN:CAA97680.1; PID:e245

A/Experimental source: strain S288C
R/Brewster, J.L.; de Valoir, T.; Dwyer, N.D.; Winter, B.; Gustin, M.C.
Science 259, 1760-1763, 1993
A/Title: An osmosensing signal transduction pathway in yeast.
A/Reference number: S42145; MIMD:93206121; PMID:7681220
A/Accession: S42145
A/Status: nucleic acid sequence not shown

A/Molecule type: DNA
A/Residues: 1-408 'GORGSCCK' <BRE>
A/Cross-references: EMBL:L06279; NID:g295610; PIDN:AAA34680.1; PID:g295611
R/Verhaaselt, P.; Volckaert, G.
submitted to the EMBL Data Library, September 1995
A/Reference number: S69393
A/Accession: S69400
A/Molecule type: DNA
A/Residues: 1-435 <VER>
A/Cross-references: EMBL:X89514; NID:g1297019; PIDN:CAA61691.1; PID:e198746; PID:g129702;
C/Genetics:
A/Gene: SGD:HOG1; SSK3
A/Cross-references: SGD:S0004103; MIPS:YLR113w
A/Map position: 12R
C/Superfamily: protein kinase homology
C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase; signal tr
F/21-271/Domain: protein kinase homology <KIN>
F/29-37/Region: protein kinase ATP-binding motif
F/144/Active site: Asp #status predicted

Query Match 10.5%; Score 231; DB 2; Length 435;
Best Local Similarity 25.5%; Pred. No. 2.5e-06;
Matches 97; Conservative 54; Mismatches 140; Indels 90; Gaps 18;

Qy 17 EAFGRLMQVQSR-----LSGSSASVYVRCCGNP--GSPGALROFLPPGTGAASA 68
Db 11 QIFGVYFETINRYNDLNPVGMGAFGLV---CSAIDTLTSGVAILKKIMKPESTAVLAQR 66
Qy 69 AEYGRKERAALEQGHENIYLVGVFTIHSPNVPSRCLLELLDVSELLSYSHQ 128
Db 67 T-----YRELKILKHLR-HENILICLDIF---LSP-LIEDYFVTELGDTLHRLT---QTR 114
Qy 129 GCSMMVIOHCARDVLEALAFHHEGYVHADLKPRNIIWSENCFKIDFGLSFKEGNQD 188
Db 115 PLEKQFQVFLQILRGKLYVASGIVHRLDIPSNILN-ENCDDKICDFGLARIQDPQM 173
Qy 189 VKYIQDGYRAPAEALQNCIAQAGLQSDTECTSAVDLWSLIGIILLEMFSGMLK---KAT 244
Db 174 TGYSTRYRAPELML-----TWQKTYVE---VDIMSAGCIFAMIBGKFLPFGKH- 222
Qy 245 VRSQEMKANSALIDHIFAS--KAVNNAIPAYHLR----- 278
Db 223 -----VHQFSIIDLDSPPKDVINTICSENTLKFTVSLPHRDPFPSRFTVEPDA 275
Qy 279 -DLISMLHDDPSRKIPAEWALCPFFSIFPAPHLEDVM-----LPTPVLR 325
Db 276 VDLLEKMLVDPDKKITTADALAH-YSAPYHDPDEPVADAKFWMHFNADLPVDTWVR 334
Qy 326 L---NYLDDVYLGNEEYEDV 343
Db 335 MMYSETLDPFKITGGSDQITDI 355

RESULT 12

A44878
protein kinase (EC 2.7.1.37) cdk2 [validated] - goldfish
C/Species: Carassius auratus (goldfish)
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C/Accession: A44878
R/Hirai, T.; Yamashita, M.; Yoshikuni, M.; Tokumoto, T.; Kajitani, H.; Sakai, N.; Nagahama
Dev. Biol. 152, 113-120, 1992
A/Title: Isolation and characterization of goldfish cdk2, a cognate variant of the cell c
A/Reference number: A44878; MIMD:92331802; PMID:1339336
A/Accession: A44878
A/Status: preliminary
A/Molecule type: mRNA

A:Residues: 1-298 <HIR>
A:Cross-references: UNIPROT:P43450; GB:S40289; NID:g251619; PIDN:AA822550.1; PID:g251622
A:Experimental source: oocyte
A:Note: sequence extracted from NCBI backbone (NCBIN:108782, NCBI:P.108783)
C:Genetics:
A:Gene: cdk2
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: ATP, cell cycle control; mitosis; phosphoprotein; phosphotransferase; serine
P:2-355/Domain: protein kinase homology <KIN>
P:10-18/Region: protein kinase ATP-binding motif
P:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 10.4%; Score 230; DB 1; Length 298;
Best Local Similarity 28.4%; Pred. No. 1.9e-06;
Matches 98; Conservative 49; Mismatches 116; Indels 82; Gaps 17;

16 LEAFRLMNOVSRLGSSGSASVYRRCNGSPGALRQF-LPPGTTGAASAAEYGR 74
Db 1 MESF-----QXVEKIGEGYGVVYKAK-NKVTGETVALKKRLDTEEGVSTAI----- 49
75 KERALEQLOLGRNIVTLVYGVFTTHFSNPVSRCLLELVSVSELLYSHQSCSMW 134
Db 50 REISLKKEL-NHNPVYKLDV--IHTENK---YVFEFLQDLKRFNDSSTVYGISLPL 103
135 IQHCARDVLEALAFIHHGGYVADIKPRNIIISAENECEFKLIDPLGSFEGNQDVKY-- 191
104 VKSYLFOLLOGLAFCHSHRVLRDLKPQNLINAGGE-IKLADGLAARAFGVPRVYTHE 162
192 IQTDSYRAREALQNCCLNAGLQSDTECTSAVDLMSIGIITLLMFPSGKMLGHTVRSQEWK 251
Db 163 VVTLVYRAREP-LIGC-----KYSTAVDIIISLCITFAEMIT-----RKALFP 204
252 ANSSAIIDHIFASKAVN-----AALPAYHL-----RDL 280
205 GDSE--IDQLFRIFPTLGTDPESIWPGVTSMPDYKSPFKMARQDLKVPPLDDEGRDL 262
281 IKSMLHDDPSRRIPAEMLCSPFSPFIPAPRIHEDLVMLPTVVL 325
Db 263 LGQMLTYPNKRKISAKNALVHRF-----RDVTM-PVPLRL 298

RESULT 13
A41227
N:Alternat names: Egl homolog; protein kinase p34
C:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text_change 09-Jul-2004
R:NhomIya-Tsuji, J.; Nomoto, S.; Yasuda, H.; Reed, S.I.; Matsumoto, K.
Proc. Natl. Acad. Sci. U.S.A. 88, 9006-9010, 1991
A:Title: Cloning of a human cDNA encoding a CDC2-related kinase by complementation of a
A:Reference number: A41227; MUID:92020980; PMID:1717994
A:Accession: A41227
A:Molecule type: mRNA
A:Residues: 1-298 <NIN>
A:Cross-references: UNIPROT:P24941; GB:M68520; NID:g180177; PIDN:AA355667.1; PID:g180178
R:Tsai, L.H.; Harlow, E.; Meyerson, M.
Nature 353, 174-177, 1991
A:Title: Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1A-
A:Reference number: S17873; MUID:91367262; PMID:1653904
A:Accession: S17873
A:Molecule type: preliminary
A:Status: preliminary
A:Residues: 1-298 <TSA>
A:Cross-references: GB:X62071; NID:g312802; PIDN:CAA43985.1; PID:g312803
R:Elledge, S.J.; Spottswood, M.R.
EMBO J. 10, 2653-2659, 1991
A:Title: A new human p34 protein kinase, CDK2, identified by complementation of a cdc28
A:Reference number: S16520; MUID:91330891; PMID:1713386
A:Accession: S16520
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-176, 'S', 178-298 <ELL>

C:Genetics: [EMBL:X61622](#); [NID:g29848](#); [PIDN:CAA43807.1](#); [PID:g29849](#)
 A:Cross-references: [GDB:128984](#); [OMIM:116953](#)
 A:Gene: [GDB:CDK2](#)
 A:Cross-references: [GDB:128984](#); [OMIM:116953](#)
 A:Map position: [12q13-12q13](#)
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: ATP, cell cycle control, mitosis; phosphoprotein; phosphotransferase; serine/threonine
 F:2-255/Domain: protein kinase homology <KIN>
 F:10-18/Region: protein kinase ATP-binding motif
 F:14-160/Binding site: phosphate (Thr) (covalent) #status predicted
 F:15/Binding site: phosphate (Tyr) (covalent) #status predicted
 F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 10.4%; Score 230; DB 2; Length 298;
 Best Local Similarity 28.7%; Pred. No. 1.9e-06;
 Matches 97; Conservative 44; Mismatches 119; Indels 78; Gaps 16;

| | | | |
|----|-----|--|-----|
| QY | 23 | WQVSRLLSGSSASYRRVCCGNGSPFGALRPF-LPPTGTAAASAEYGFRRERALE | 81 |
| DB | 4 | FOVKERKIGEGYGVVYKAR--NKLTEGVVALKKIRIDTEEGVSTNI-----REISLKL | 56 |
| QY | 82 | QLQGRNITYTLGVFTFHSPVNSRCLLELELDVSEVSELLVSSHOGSCMMNIQHCARD | 141 |
| DB | 57 | EL-NHPNIYKLDV--IHTENKL--YLVPEFLHODLKKEMDSALTGIFPLPIKSYLQ | 110 |
| QY | 142 | VLEALAPLHHEGYVHADLKPRNIIWSAENECFKLIDFGISFEKGNQDVKY--IQTDGYR | 198 |
| DB | 111 | LLQGLAFCHSHRVLRHDLKPQMLINTEG-AIKLDAFGILARAGVPRVRYTHEVVTLMYR | 169 |
| QY | 199 | APREALQNCIAAGIQSDTECTSAVDLWSIGITLLMFSGMKLKHVTSQEWKANSAIL | 258 |
| DB | 170 | APET-LLGQ-----KYSTAVDWSLGLCTFAEM-----VTRRALFEDSE--I | 209 |
| QY | 259 | DHIFASKAVVN-----AAIPAYHL-----RDILKSMILMD | 287 |
| DB | 210 | DQLFRIFRILGTPDEVVMPGVGTSMPDYKPSFKMARQDFSKVVPPLDEDGRSLISQMLHY | 269 |
| QY | 288 | DPSRRIPAEMLCSPPFSSIFPAFHIEDLVMLPVPVRL | 325 |
| DB | 270 | DPNKRIISAKAALAHPPF-----QD-VTKVPVPHRL | 298 |

RESULT 14
 S23387
 protein kinase (EC 2.7.1.37) cdK6 - human
 N:Alternate names: cyclin-dependent kinase 6; protein kinase cdK2-related PLSTIRE
 C:Species: Homo sapiens (man)
 C:Dates: 18-Jun-1993 #sequence_revision 18-Jun-1993 #text_change 09-Jul-2004
 C:Accession: [S23387](#); [S22749](#)
 R:Mayerson, M.; Ender, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; Teasdale
 A>Title: A family of human cdK2-related protein kinases.
 A:Accession: [S23387](#)
 A:Accession: [S23387](#); [MUID:92347325](#); [PMID:1539063](#)
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-326 <MEV>
 A:Cross-references: [UNIPROT:Q00534](#); [EMBL:X66365](#); [NID:g36622](#); [PIDN:CAA47008.1](#); [PID:g36623](#)
 C:Genetics:
 A:Gene: [GDB:CDK6](#)
 A:Cross-references: [GDB:283458](#)
 A:Map position: [7q36-7q36](#)
 C:Superfamily: kinase-related transforming protein; protein kinase homology
 C:Keywords: ATP, cell cycle control, phosphoprotein; phosphotransferase; serine/threonine
 F:11-284/Domain: protein kinase homology <KIN>
 F:19-27/Region: protein kinase ATP-binding motif
 F:24/Binding site: phosphate (Tyr) (covalent) #status predicted
 F:43,61,145,147/Active site: Lys, Glu, Asp, Lys #status predicted
 F:177/Binding site: phosphate (Thr) (covalent) #status predicted
 F:299/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 10.4%; Score 230; DB 2; Length 326;
 Best Local Similarity 27.4%; Pred. No. 2.1e-06;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: November 29, 2004, 11:26:47 ; Search time 122.5 Seconds
(without alignments)
1968.016 Million cell updates/sec

Title: US-10-798-532-4

Perfect score: 2206
Sequence: 1 MAGSGCAGCAEPPRFLEARG.....VATFVPLSAVKKGYLYQTLL 419

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 2195 | 99.5 | 419 | 1 | KIST_HUMAN |
| 2 | 2183 | 99.0 | 419 | 1 | KIST_RAT |
| 3 | 2182 | 98.9 | 419 | 2 | AABH5732 |
| 4 | 2182 | 98.9 | 419 | 2 | AAOI3515 |
| 5 | 2177 | 98.7 | 419 | 1 | KIST_MOUSE |
| 6 | 1711 | 77.6 | 330 | 2 | Q8CON6 |
| 7 | 254 | 11.5 | 505 | 2 | Q6CFX3 |
| 8 | 248.5 | 11.3 | 882 | 2 | Q8MOL7 |
| 9 | 242 | 11.0 | 433 | 2 | Q84SN3 |
| 10 | 240.5 | 10.9 | 395 | 2 | Q6GM90 |
| 11 | 240.5 | 10.9 | 720 | 2 | Q9CA22 |
| 12 | 240.5 | 10.9 | 765 | 2 | Q25758 |
| 13 | 240.5 | 10.9 | 777 | 2 | Q84VX4 |
| 14 | 240.5 | 10.9 | 826 | 2 | Q94656 |
| 15 | 240.5 | 10.9 | 914 | 2 | Q8LIF0 |
| 16 | 240 | 10.9 | 368 | 1 | SMK1_YEAST |
| 17 | 239.5 | 10.9 | 395 | 2 | Q8AYR3 |
| 18 | 239.5 | 10.9 | 395 | 1 | AABH7045 |
| 19 | 238.5 | 10.8 | 387 | 2 | SKP1_SCHPO |
| 20 | 237.5 | 10.8 | 360 | 2 | Q6S677 |
| 21 | 237.5 | 10.8 | 360 | 2 | AAR17088 |
| 22 | 235.5 | 10.7 | 434 | 2 | Q75B85 |
| 23 | 235.5 | 10.7 | 434 | 2 | AAS51605 |
| 24 | 235 | 10.7 | 298 | 1 | CDK2_MESAU |
| 25 | 234.5 | 10.6 | 392 | 2 | Q6ZD93 |
| 26 | 234.5 | 10.6 | 392 | 2 | BAC9508 |
| 27 | 234.5 | 10.6 | 433 | 1 | KKO1_YEAST |
| 28 | 234 | 10.6 | 376 | 1 | MPK4_ARATH |
| 29 | 233.5 | 10.6 | 433 | 2 | Q6B829 |
| 30 | 233.5 | 10.6 | 433 | 2 | Q6B829 |
| 31 | 233.5 | 10.6 | 1233 | 2 | Q6CAC2 |

| | | | | | |
|----|-------|------|-----|---|------------|
| 32 | 233 | 10.6 | 369 | 2 | Q7M445 |
| 33 | 232.5 | 10.5 | 367 | 1 | MK12_RAT |
| 34 | 232 | 10.5 | 568 | 2 | Q9HBV6 |
| 35 | 232 | 10.5 | 588 | 2 | Q9HBY7 |
| 36 | 232 | 10.5 | 588 | 2 | AAT06103 |
| 37 | 231.5 | 10.5 | 367 | 1 | MK12_MOUSE |
| 38 | 231.5 | 10.5 | 464 | 1 | MK10_RAT |
| 39 | 231.5 | 10.5 | 542 | 2 | Q61587 |
| 40 | 231 | 10.5 | 298 | 1 | CDK2_CRIGR |
| 41 | 231 | 10.5 | 298 | 1 | CDK2_RAT |
| 42 | 231 | 10.5 | 298 | 2 | Q6F751 |
| 43 | 231 | 10.5 | 298 | 2 | AAH61832 |
| 44 | 231 | 10.5 | 361 | 1 | JNK_SUDDO |
| 45 | 231 | 10.5 | 435 | 1 | HOG1_YEAST |

ALIGNMENTS

RESULT 1
KIST_HUMAN
ID KIST_HUMAN STANDARD; PRT; 419 AA.
AC Q8TAS1; Q96C22;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting with actin).
GN Name=KIST; Synonyms=KIS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX PubMed=12782393;
RA Bieche I., Manceau V., Curni P.A., Laurendeau I., Lachkar S., Leroy K., Vidard D., Sobel A., Maucuer A.;
RT "Quantitative RT-PCR reveals a ubiquitous but preferentially neural expression of the KIS gene in rat and human."
RL Brain Res. Mol. Brain Res. 114:55-64(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Eye, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.D., Ueda T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malak J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Possesses kinase activity. May be involved in trafficking and/or processing of RNA (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with actin (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;

```

CC      IsoId=Q8TRA1-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=Q8TRA1-2; Sequence=VSP_004908, VSP_004909;
CC      Note=No experimental confirmation available;
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC      -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; A1536197; CAD60192.1; -
DR      EMBL; BC014917; AAH14917.1; -
DR      EMBL; BC026046; AAH26046.1; -
DR      HSSP; P26368; 100P.
DR      GO; GO:0005737; C:cytoplasm; ISS.
DR      GO; GO:0005634; C:nucleus; ISS.
DR      GO; GO:0005524; F:ATP binding; ISS.
DR      GO; GO:0005515; F:protein binding; ISS.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR      GO; GO:0003723; F:RNA binding; ISS.
DR      GO; GO:0018105; F:peptidyl-serine phosphorylation; ISS.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR000719; Prot kinase.
DR      InterPro; IPR000504; RNA_rec_mot.
DR      InterPro; IPR008271; Ser_thr_pkin_AS.
DR      Pfam; PF00069; Pkinase; 1.
DR      Pfam; PF00076; RNM_1; 1.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      SMART; SM00360; RNM; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR      PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR      PROSITE; PS00102; RNM; 1.
DR      PROSITE; PS00030; RNM_RNP_1; FALSE_NEG.
KW      Alternative splicing; ATP-binding; Nuclear protein; RNA-binding;
KW      Serine/threonine-protein kinase; Transferase.
FT      DOMAIN 23..304
FT      NP_BIND 324..406
FT      BINDING 29..37
FT      ACT_SITE 54..58
FT      VARSPLIC 342..344
FT      VARSPLIC 345
FT      VARSPIC 419
SQ      SEQUENCE 419 AA; 46546 MW; 903E982BE12A8CF8 CRC64;
Query Match 99.5%; Score 2195; DB 1; Length 419;
Best Local Similarity 99.5%; Pred. No. 2.4e-158;
Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

DB      241 LKHTVRSQEWKANSALIDHIFASKAVVNAALPAYHLRLIKSMHDDPSRRIPAEMLC 300
QY      301 SPFSIPFAPHIEDLVMLPTPLRLNVLDDDYLCNESEYEDVVDVKECKQKPYVSL 360
DB      301 SPFSIPFAPHIEDLVMLPTPLRLNVLDDDYLCNESEYEDVVDVKECKQKPYVSL 360
QY      361 LVPKENPGRGQFVEVYANAGDSKAAQKLTGRMPDGKVVATFYPPLSAVKGQYQTL 419
DB      361 LVPKENPGRGQFVEVYANAGDSKAAQKLTGRMPDGKVVATFYPPLSAVKGQYQTL 419
RESULT 2
ID      KIST RAT          STANDARD;          PRT;          419 AA.
AC      063285;
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Serine/threonine-protein kinase Kist (EC 2.7.1.37) (kinase interacting
DE      with strachin) (PAM COOH-terminal interactor protein 2) (P-CIP2).
GN      Name=Kist; Synonyms=Kis;
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxId=10116;
RN      [1]
RP      SEQUENCE FROM N.A., FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, TISSUE
RP      SPECIFICITY, AND MUTAGENESIS OF LYS-54.
RC      TISSUE=striatum;
RX      MEDLINE=97435279; PubMed=9287318;
RA      Maucuer A., Ozon S., Mancieu V., Gavet O., Lawler S., Curmi P.,
RA      Sobel A.;
RT      "Kis is a protein kinase with an RNA recognition motif.";
RL      J. Biol. Chem. 272:23151-23156(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Caldwell B.D., Darlington D.N., Penzes P., Johnson R.C., Eipper B.A.,
RA      Mains R.E.;
RL      Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE OF 28-419 FROM N.A.
RC      TISSUE=hippocampus;
RX      MEDLINE=97067094; PubMed=8910496;
RA      Alam R., Caldwell B.D., Johnson R.C., Darlington D.N., Mains R.E.,
RA      Eipper B.A.;
RT      "Novel proteins that interact with the COOH-terminal cytosolic routing
RT      determinants of an integral membrane peptide-processing enzyme.";
RL      J. Biol. Chem. 271:28636-28640(1996).
CC      -1- FUNCTION: Possesses kinase activity. May be involved in
CC      trafficking and/or processing of RNA.
CC      -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC      -1- SUBUNIT: Interacts with strachin (By similarity).
CC      -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC      -1- TISSUE SPECIFICITY: In the embryo, preferentially expressed in the
CC      developing nervous system.
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC      -1- SIMILARITY: Contains 1 RNA recognition motif (RNM) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X98374; CAA67021.1; -
DR      EMBL; U70372; AAC53031.2; -
DR      HSSP; P26368; 100P.
DR      RGD; 2968; Kist.
DR      GO; GO:0005737; C:cytoplasm; IDA.
DR      GO; GO:0005634; C:nucleus; IDA.

```

DR GO; GO:0005524; F:ATP binding; IC.
 DR GO; GO:0005515; F:protein binding; ISS.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
 DR GO; GO:0003723; F:RNA binding; NAS.
 DR GO; GO:0018105; F:peptidyl-serine phosphorylation; IDA.
 DR InterPro; IPR011009; Kinase_like.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR005054; RNA_rec_mec.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; I.
 DR Pfam; PF00076; RM_1; 1.
 DR ProDom; PD000001; Prot_Kinase; 1.
 DR SMART; SM00360; RM; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE; PS50102; RM; 1.
 DR PROSITE; PS00030; RM_RNF_1; FALSE_NEG.
 DR ATP-binding; Nuclear protein; RNA-binding;
 KW Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 23 304 Protein kinase.
 FT NP_BIND 324 406 RNA-binding (RM).
 FT BINDING 29 37 ATP (By similarity).
 FT ACT_SITE 54 54 ATP.
 FT ACT_SITE 141 141 Proton acceptor (By similarity).
 FT ACT_SITE 158 158 Proton acceptor (By similarity).
 FT MUTAGEN 54 54 K->R: Loss of activity.
 SQ SEQUENCE 419 AA; 46505 MW; D2F87BF8P080DF3 CRC64;

Query Match 99.0%; Score 2183; DB 1; Length 419;
 Best Local Similarity 98.8%; Pred. No. 1.9e-157;
 Matches 414; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGSGCAGAAEPRLLEAFGRMLQVQSRIGSSASVYRVCCGPGSPGALROFLPPG 60
 DB 1 MAGSGCAGAAEPRLLEAFGRMLQVQSRIGSSASVYRVCCGPGSPGALROFLPPG 60
 QY 61 TTGAASAABEYGFRRKRALBOLQGHRIYVLYGVFTIHFSFNVPSRCLLELDVSVSE 120
 DB 61 TTGAASAABEYGFRRKRALBOLQGHRIYVLYGVFTIHFSFNVPSRCLLELDVSVSE 120
 QY 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLEHGYHADLKPRNIIWASAENECFKLIDFGL 180
 DB 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLEHGYHADLKPRNIIWASAENECFKLIDFGL 180
 QY 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLEHGYHADLKPRNIIWASAENECFKLIDFGL 180
 DB 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLEHGYHADLKPRNIIWASAENECFKLIDFGL 180
 QY 181 SREKGNQDVYKTYQTDGYRAPEALONCLAQGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
 DB 181 SREKGNQDVYKTYQTDGYRAPEALONCLAQGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
 QY 241 LKHTVRSQEWKANSALIDHIFASKAVVNAAPAYHLRDLIKSMHDDPSRRIPEAMALC 300
 DB 241 LKHTVRSQEWKANSALIDHIFASKAVVNAAPAYHLRDLIKSMHDDPSRRIPEAMALC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDVLTGNEEYEDVDEVDKECKYGPVSL 360
 DB 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDVLTGNEEYEDVDEVDKECKYGPVSL 360
 QY 361 LVPKENPGRGVFEVYANAGDSKAAOKLLTGMPFGKFVVAFTFPLSAVKGYLYQTLL 419
 DB 361 LVPKENPGRGVFEVYANAGDSKAAOKLLTGMPFGKFVVAFTFPLSAVKGYLYQTLL 419

RESULT 3
 AAH58732
 ID AAH58732 PRELIMINARY; PRT; 419 AA.
 AC AAH58732;
 DT 02-MAR-2004 (Tremblrel. 27, Created)
 DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
 DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)
 DE Kinase interacting with leukemia-associated gene (Scatmin).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OK NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Szeitz T.E.,
 RA Brownstein M.J., Umed T.B., Toshiki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kertman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RP [2]
 RC SEQUENCE FROM N.A.
 RA STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.,
 RL Submitted (Sep-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC058732; AAH58732.1; -.
 KW Kinase.
 SQ SEQUENCE 419 AA; 46489 MW; 3BD3C06A59E22540 CRC64;

Query Match 98.9%; Score 2182; DB 2; Length 419;
 Best Local Similarity 98.8%; Pred. No. 2.3e-157;
 Matches 414; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAGSGCAGAAEPRLLEAFGRMLQVQSRIGSSASVYRVCCGPGSPGALROFLPPG 60
 DB 1 MAGSGCAGAAEPRLLEAFGRMLQVQSRIGSSASVYRVCCGPGSPGALROFLPPG 60
 QY 61 TTGAASAABEYGFRRKRALBOLQGHRIYVLYGVFTIHFSFNVPSRCLLELDVSVSE 120
 DB 61 TTGAASAABEYGFRRKRALBOLQGHRIYVLYGVFTIHFSFNVPSRCLLELDVSVSE 120
 QY 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLEHGYHADLKPRNIIWASAENECFKLIDFGL 180
 DB 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLEHGYHADLKPRNIIWASAENECFKLIDFGL 180
 QY 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLEHGYHADLKPRNIIWASAENECFKLIDFGL 180
 DB 121 LLLYSSHGSCMMWMIQHCAVDYLEALAFLEHGYHADLKPRNIIWASAENECFKLIDFGL 180
 QY 181 SREKGNQDVYKTYQTDGYRAPEALONCLAQGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
 DB 181 SREKGNQDVYKTYQTDGYRAPEALONCLAQGLQSDTECTSAVDLMSLGIILLEMFSGMK 240
 QY 241 LKHTVRSQEWKANSALIDHIFASKAVVNAAPAYHLRDLIKSMHDDPSRRIPEAMALC 300
 DB 241 LKHTVRSQEWKANSALIDHIFASKAVVNAAPAYHLRDLIKSMHDDPSRRIPEAMALC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDVLTGNEEYEDVDEVDKECKYGPVSL 360
 DB 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDVLTGNEEYEDVDEVDKECKYGPVSL 360
 QY 361 LVPKENPGRGVFEVYANAGDSKAAOKLLTGMPFGKFVVAFTFPLSAVKGYLYQTLL 419
 DB 361 LVPKENPGRGVFEVYANAGDSKAAOKLLTGMPFGKFVVAFTFPLSAVKGYLYQTLL 419

RESULT 4
 AA013515
 ID AA013515 PRELIMINARY; PRT; 419 AA.
 AC AA013515;
 DT 02-MAR-2004 (Tremblrel. 27, Created)
 DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)

```

DT 02-MAR-2004 (TreeMrel. 27, last annotation update)
DE KIS kinase.
GN KIS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=12956;
RA Crook M.F., Boehm M., Nabel E.G.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY180177; AAO13515.1; -.
KW Kinasase.
SQ SEQUENCE 419 AA; 46489 MW; 3BD3C06A59E2540 CRC64;

Query Match 98.9%; Score 2182; DB 2; Length 419;
Best Local Similarity 98.8%; Pred. No. 2.3e-157;
Matches 414; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAGSGCAGMGAEPFRPLFAFGRLMVOVSRLGSSASAYRVRCGNGSPGALROFLPPG 60
DB 1 MAGSGCAGMGAEPFRPLFAFGRLMVOVSRLGSSASAYRVRCGNGSPGALROFLPPG 60
QY 61 TTGAASAAYEYGRKRAALQOQHNRNIVLYGVFTIHSPVPSRCILLLELDVSVS 120
DB 61 TTGAASAAYEYGRKRAALQOQHNRNIVLYGVFTIHSPVPSRCILLLELDVSVS 120
QY 121 LLLYSHQCGSMWMIQHCAVDVLEALFLHGEVYADLKPRYIWSAENECFKLIDFGL 180
DB 121 LLLYSHQCGSMWMIQHCAVDVLEALFLHGEVYADLKPRYIWSAENECFKLIDFGL 180
QY 181 SFKEGNQDVYIOTDGYRAPEALQNCIAQGLQSDTECTSAVDLSLGIILLEMESGK 240
DB 181 SFKEGNQDVYIOTDGYRAPEALQNCIAQGLQSDTECTSAVDLSLGIILLEMESGK 240
QY 241 LKHTVRSQEWKANSIAIITHIPASKAVNNAIPAYHLRLIKSLHDDPSRRIPAEALC 300
DB 241 LKHTVRSQEWKANSIAIITHIPASKAVNNAIPAYHLRLIKSLHDDPSRRIPAEALC 300
QY 301 SPFFSIPFAHIDLVLPVRLNVLDLDDYLENDEDEVEDVEDEKCEQKIGVSVSL 360
DB 301 SPFFSIPFAHIDLVLPVRLNVLDLDDYLENDEDEVEDVEDEKCEQKIGVSVSL 360
QY 361 LVPKEMPGQGVVEYANAGDSKAAQKLTLMFMFGKVVATFPYLSAYKRGYLTLL 419
DB 361 LVPKEMPGQGVVEYANAGDSKAAQKLTLMFMFGKVVATFPYLSAYKRGYLTLL 419
DB (1)
RESULT 5
KIST MOUSE STANDARD; PRT; 419 AA.
ID KIST_MOUSE
AC P97313; O61775; O9CYT1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase Kist (EC 2.7.1.37) (Kinase interacting
DE with strathmin).
GN Name=Kist; Synonym=KIS;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
MEDLINE=22354663; PubMed=1246851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Oosato N., Saito R., Suzuki A., Yamanaka I., Kiyosawa H.,
RA Nagai K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schirml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Brand D., Brusic V., Chochia C., Corbani L.E., Cousins S.,

```

```

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Nunata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Maki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
RN (2)
RP SEQUENCE OF 1-414 FROM N.A.
RC TISSUE=Brain;
MEDLINE=97435279; PubMed=9287318;
RA Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P.,
RA Sobel A.;
RT "KIS is a protein kinase with an RNA recognition motif."
RL J. Biol. Chem. 272:23151-23156 (1997).
RN (3)
RP REVISIONS TO 52 AND 68-69.
RA Maucuer A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE OF 122-321 FROM N.A., AND INTERACTION WITH STRATHMIN.
RC TISSUE=Embryo;
MEDLINE=95241452; PubMed=7724523;
RA Maucuer A., Camonis J.H., Sobel A.;
RT "Strathmin interaction with a putative kinase and coiled-coil-forming
RT protein domains."
RL Proc. Natl. Acad. Sci. U.S.A. 92:3100-3104 (1995).
CC -1- FUNCTION: Possesses kinase activity. May be involved in
CC trafficking and/or processing of RNA (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated
CC -1- SUBUNIT: Interacts with strathmin.
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-ch).
CC -----
DR EMBL: Y10725; CAA1714.2; -.
DR EMBL: AK01347; BAB28802.1; -.
DR EMBL: X82320; CAA57763.1; -.
DR PIR: I48615; I48615.
DR HSSP: P26368; 100P.
DR MGD: MGI:1341908; Kist.
DR GO: GO:0005737; C:cytoplasm; ISS.
DR GO: GO:0005634; C:nucleus; ISS.
DR GO: GO:0005524; F:ATP binding; ISS.
DR GO: GO:0005515; F:protein binding; IPI.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; ISS.
DR GO: GO:0003723; F:RNA binding; ISS.
DR GO: GO:0018105; P:peptidyl-serine phosphorylation; ISS.
DR InterPro: IPR011009; Kinase_like.

```

DR InterPro: IPR000719; Prot kinase.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR InterPro: IPR008271; Ser_thr_pkin_AS.
 DR Pfam: PF00069; Pkinase; 1.
 DR Pfam: PF00076; RRM_1; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00360; RRM; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE: PS50102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 DR ATP-binding; Nuclear protein; RNA-binding;
 KM Serine/threonine-protein kinase; Transferase.
 FT DOMAIN 23 303
 FT NE_BIND 323 405 RNA-binding (RRM).
 FT BINDING 29 37 ATP (By similarity).
 FT ACT_SITE 158 158 ATP (By similarity).
 FT CONFLICT 170 170 Proton acceptor (By similarity).
 FT CONFLICT 231 231 N -> D (in Ref. 4).
 SQ SEQUENCE 419 AA; 46477 MW; A91B307BA50F29A1 CRC64;

Query Match 98.7%; Score 2177; DB 1; Length 419;
 Best Local Similarity 98.6%; Pred. No. 5.5e-157;
 Matches 413; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MASSGCAMGAPRRFLFAFRRLNQVGRSLGSSASAYRRCCNGSPFGALROFLPPG 60
 1 MASSGCAMGAPRRFLFAFRRLNQVGRSLGSSASAYRRCCNGSPFGALROFLPPG 60

DB 61 TTGAASAAEYGRKERAALEQGHNIYTLGVFTIHSPVNPSCRLLLELDVSVSE 120
 61 TTGAASAAEYGRKERAALEQGHNIYTLGVFTIHSPVNPSCRLLLELDVSVSE 120

QY 121 LLLYSSHQGSMWMIQHCARDVLEALFLHHEGVHADLPRNIIILSAENECFKLIDFGL 180
 121 LLLYSSHQGSMWMIQHCARDVLEALFLHHEGVHADLPRNIIILSAENECFKLIDFGL 180

DB 121 LLLYSSHQGSMWMIQHCARDVLEALFLHHEGVHADLPRNIIILSAENECFKLIDFGL 180
 121 LLLYSSHQGSMWMIQHCARDVLEALFLHHEGVHADLPRNIIILSAENECFKLIDFGL 180

QY 181 SFEGKQNDVYKITTQDGRAPAEALONCLAQGLQSDTECTSAVDLWSLGLTLEMFSGMK 240
 181 SFEGKQNDVYKITTQDGRAPAEALONCLAQGLQSDTECTSAVDLWSLGLTLEMFSGMK 240

DB 181 SFEGKQNDVYKITTQDGRAPAEALONCLAQGLQSDTECTSAVDLWSLGLTLEMFSGMK 240
 181 SFEGKQNDVYKITTQDGRAPAEALONCLAQGLQSDTECTSAVDLWSLGLTLEMFSGMK 240

QY 241 LKHTVRSQEWKANSATIDHIFASKAVVNAIPAYHLRDLIKSLMHDDESRRIIPAEALC 300
 241 LKHTVRSQEWKANSATIDHIFASKAVVNAIPAYHLRDLIKSLMHDDESRRIIPAEALC 300

DB 241 LKHTVRSQEWKANSATIDHIFASKAVVNAIPAYHLRDLIKSLMHDDESRRIIPAEALC 300
 241 LKHTVRSQEWKANSATIDHIFASKAVVNAIPAYHLRDLIKSLMHDDESRRIIPAEALC 300

QY 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVLDVLYGNEEYEDVVDYKECQKYGPVSL 360
 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVLDVLYGNEEYEDVVDYKECQKYGPVSL 360

DB 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVLDVLYGNEEYEDVVDYKECQKYGPVSL 360
 301 SPFFSIPFAPIHEDLVMLPTPVRLNLNVLDVLYGNEEYEDVVDYKECQKYGPVSL 360

QY 361 LVKENPGRQVVEVEYANAGDSKAAQKLLTGRNFDGKFVATYTPISAYKRGYLYOTLL 419
 361 LVKENPGRQVVEVEYANAGDSKAAQKLLTGRNFDGKFVATYTPISAYKRGYLYOTLL 419

DB 361 LVKENPGRQVVEVEYANAGDSKAAQKLLTGRNFDGKFVATYTPISAYKRGYLYOTLL 419
 361 LVKENPGRQVVEVEYANAGDSKAAQKLLTGRNFDGKFVATYTPISAYKRGYLYOTLL 419

RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Mech. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN PANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA The PANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carninci P., Kono H., Akiyama Y., Nishi K., Katsunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kanakawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohse N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasai D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Tota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK030152; BAC26812.1; -.
 DR HSSP; P26368; 100P.
 DR MGD; MGI:1341908; Kiat.
 DR GO; GO:0005524; F-ATP binding; IEA.
 DR GO; GO:0016740; F-transportase activity; IEA.
 DR GO; GO:0004468; P-protein amino acid phosphorylation; IEA.
 DR InterPro: IPR001009; Kinase-like.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00069; Pkinase; 2.
 DR Pfam: PF00076; RRM_1; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

| | |
|----------|---|
| DR | PROSITE; PS50102; RRM; 1. |
| KW | kinase. |
| SO | SEQUENCE 330 AA; 37201 MW; 4164C0758BE7CB4C CRC64; |
| OY | Query Match 77.6%; Score 1711; DB 2; Length 330; Best Local Similarity 98.8%; Pred. No.1.le-121; Matches 326; Conservative 2; Mismatches 2; Indels 0; Gaps 0; |
| Dy | 90 VTLGVPTHTSPNPNRSCLLELTDVSVSELTLYSSHQSGSMWMIOHCARDVLEALFL 149 ::: |
| Dd | 1 MTLGVFTTHSPNPVPSRCLLELTDVSVSELTLYSSHQSGSMWMIOHCADVDLEALFL 60 |
| Oy | 150 HHEGVHADLPKRNILMSAENECFKLIDFGSFKEGNDVXYIQTGYRAPEALONCIA 209 |
| Dd | 61 HHEGVHADLPKRNILMSAENECFKLIDFGSFKEGNDVXYIQTGYRAPEALONCIA 120 |
| Oy | 210 OAGQSOPTECHSAVDLMSLGILLMEFSGMKLKHVRQEWKANSALIIDHIFASKAVVN 269 |
| Dd | 121 QAGQSODETCSAVDLMSLGILLMEFSGMKLKHVRQEWKANSALIIDHIFASKAVVN 180 |
| Oy | 270 AAIAYHLRDLIKSMHDDPERRIPAEANALCSFFESIFPAHIEDLVMLPTVLRLLNLV 329 |
| Dd | 181 AAIAYHLRDLIKSMHDDPERRIPAEANALCSFFESIFPAHIEDLVMLPTVLRLLNLV 240 |
| Oy | 330 DDVYLGNDEEVEDVVEKYECOKTPVVSLLVPKNPGRGVFPYEYNAGDSKAQQL 389 |
| Dd | 241 DDVYLENDEVEDVVEKYECOKTPVVSLLVPKNNGRGVFPVEYNAGDSKAQQL 300 |
| Oy | 390 TGRMPDGKFVVATFPPLSAVKRYGYOTLL 419 |
| Dd | 301 TGRMPDGKFVVATFPPLSAVKRYGYOTLL 330 |
| RESULT 7 | |
| ID | 06CFX3 PRELIMINARY; PRT; 505 AA. |
| AC | 06CFX3; |
| DT | 01-OCT-2004 (TREMBLrel. 28, Created) |
| DT | 01-OCT-2004 (TREMBLrel. 28, Last sequence update) |
| DT | 01-OCT-2004 (TREMBLrel. 28, Last annotation update) |
| DE | Chromosome B of strain CLIB99 of Yarrowia lipolytica. ORNames=YALI0B02816g; |
| GN | Yarrowia lipolytica (Candida lipolytica). |
| OC | Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; |
| CC | Saccharomycetales; Dipodascaceae; Yarrowia. |
| OX | NCHI_TaxID=4952; |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN=CLIB99; |
| RG | GENOLEVURES; |
| RA | Lafontaine I., de Montigny J., Matck C., Neuvéglise C., Talia E., Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S., Ra Goiffard N., Frangeul S., Aigle M., Anhouard V., Babour A., Barbe V., Batney S., Bianchini S., Beckerich J.M., Beyne E., Bleykasten C., Boisarme A., Boyer J., Catolico L., Confanieret F., de Darvar A., Depons L., Fabre E., Fairhead C., Ferry-Dunazet H., Groppi A., Hartveit F., Hennequin C., Jauniaux N., Joyet P., Kachouri R., Kerrest A., Koszul R., Lemaitre M., Iseur I., Ma L., Muller H., Micaud J.M., Nikolski M., Oztes S., Ozier-Kalogeropoulos O., Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A., Swenne D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B., Zentou-Meyer M., Zivanovic I., Bojotin-Fukuhara M., Thierry A., Ra Bouchier C., Caubron B., Scarpell C., Gallardin C., Weissbach J., Wincker P., Souciet J.L.; RT "Genome evolution in yeasts."; RL Nature 430:35-44(2004). [2] |
| RN | SEQUENCE FROM N.A. |
| RP | STRAIN=CLIB99; |
| RC | Genoscope; |
| RA | Submitted (JUL-2004) to the EMBL/Genbank/DDBJ databases. |
| EMBL | CR362128; CAG62657.1; - |
| DQ | SEQUENCE 505 AA; 58191 MW; D30933013FCBE357 CRC64; |

| Query Match | Similarity | Score | DB 2 | Length | 505 |
|-------------|------------|-------------------|------|------------|-----|
| Best Local | 25.2% | Pred. No. 9,58-11 | | | |
| Matches | 87 | Conservative | 47 | Mismatches | 129 |
| | | | | Indels | 82 |
| | | | | Gaps | 11 |

| Query | Subject | Score | DB 2 | Length | 505 |
|-------|--|-------|------|--------|-----|
| QY | W0V0SRLLGSGSSASVAVRRCGNGSPGALRQFLPPTGTGAASAAEYGRKERALEQ | 82 | | | |
| Db | F0VTEKELGHGAYGIVCAKATGTGTDEGVAIKVTNIFSKNIIKRA-----LREIKLNLH | 77 | | | |
| QY | LOGHRNIVTLXGVFTTHFSRNVPRCILLLELDV-SVSEILLYSGHSCSMMMI----- | 135 | | | |
| Db | 78 FRGHKNITCTCYD-----MDIVDTNNFELLYEELMECDHQTIRSGQ | 121 | | | |
| QY | 136 -----OHCARDVLEALFLHHEGYVHADLKPRNIIILSAENECEKLIIDFGSLFKEGNDV | 189 | | | |
| Db | 122 LITDAHYGSPFYQILAGKVIYHSADVLRDLKRGSLVNLVADDE-LKIDCGFLARGSSDDE | 180 | | | |
| QY | 190 K-----YIQTDDGYRPAPEALONCLAOAGIQSDPCTSAVDLWSLGIILLEMFSKMKL | 242 | | | |
| Db | 181 KNAGFLTEYVATRYRPAPEIML-----SFQSYTKAIDIMSVGCIILLELIGG---K | 227 | | | |
| QY | 243 HTVNSQGMKANSKSIIDHI-----FASKRVNAAIPAYH-- | 276 | | | |
| Db | 228 PLEFGKNVVDQNLILHYLTGTPSEETLKRIKSGPRAQEVYRGLPMPKIPFSTLEPTANPE | 287 | | | |
| QY | 277 LRDILKSMLEHDDPSRRIPAEWALCSPPFSIIPFAPIHEDLVMLPTP | 321 | | | |
| Db | 288 ALDLEKMLAFDPARERTVEBALHPILKIMHPD--RDEPVCPTP | 330 | | | |

| RESULT 8 | PRELIMINARY | PRT | 882 AA. |
|----------|--|---|---------|
| ID | 08WOL7 | | |
| AC | 08WOL7 | 022155; | |
| DT | 01-MAR-2002 | (TReMBLrel. 20, Created) | |
| DT | 01-MAR-2002 | (TReMBLrel. 20, Last sequence update) | |
| DT | 01-OCT-2004 | (TReMBLrel. 28, Last annotation update) | |
| DE | Man1b1rain kinase (Hypochemical protein T04C10.1). | | |
| GN | Name=mbk-1; Synonym=T04C10.1; | | |
| OS | Caenorhabditis elegans. | | |
| OC | Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides; | | |
| OC | Rhabditidae; Peloiderinae; Caenorhabditis. | | |
| OX | NCBI_TaxID=6239; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=22505392; Pubmed=12618396; | | |
| RA | Ratich W.B., Mooman C., Laceyfield C.O., Lehrer J., Bartesch D., | | |
| RA | Plasterk R.H., Kandel E.R., Hobert O.; | | |
| RT | "Characterization of Caenorhabditis elegans homologs of the Down | | |
| RT | syndrome candidate gene DYRK1A."; | | |
| RL | Genetics 163:571-580(2003). | | |
| RN | [2] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=Briscot N2; | | |
| RX | MEDLINE=99069613; Pubmed=9851916; | | |
| RA | none; | | |
| RT | "Genome sequence of the nematode C.elegans: A platform for | | |
| RT | investigating biology."; | | |
| RL | Science 282:2012-2018(1998). | | |
| RN | [3] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=Briscot N2; | | |
| RA | Burton J.; | | |
| RA | Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases. | | |
| CC | -I- SMILIARTY: Belongs to the Ser/thr protein kinase family. | | |
| DR | EMBL; AY064444; AAL40874.1; - | | |
| DR | EMBL; Z69885; CA93756.2; - | | |
| DR | HSSP; P24941; 1H0V. | | |
| DR | Interact; Q8WOL7; - | | |
| DR | GO; GO:0005524; F:ATP binding; IEA. | | |
| DR | GO; GO:0004674; F:protein serine/threonine kinase activity; IEA. | | |
| DR | GO; GO:0016740; F:transferase activity; IEA. | | |
| DR | GO; GO:0006468; P:protein amino acid phosphorylation; IEA. | | |

| | |
|-----------------------|---|
| RA | Klausner D.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., |
| RA | Aleisch S.F., Zeeberg B., Bucow K.H., Schaefer C.F., Bhat N.K., |
| RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hesleh F., |
| RA | Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., |
| RA | Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., |
| RA | Brownstein M.J., Uediri T.B., Toshiyuki S., Caranci P., Prange C., |
| RA | Raba S.S., Luquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., |
| RA | Boeak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H., |
| RA | Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., |
| RA | Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., |
| RA | Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., |
| RA | Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., |
| RA | Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., |
| RA | Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., |
| RA | Krzywinski W.I., Skalska U., Smalios D.E., Schmerch A., Schein J.E., |
| RA | Jones S.U., Marx M.A., |
| RT | "Generation and initial analysis of more than 15,000 full-length human |
| RT | and mouse cDNA sequences." |
| RL | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). |
| RP | [2] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE=Kidney; |
| RX | MEDLINE=22341132; PubMed=12454917; |
| RA | Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W., |
| RA | Richardson P.; |
| RT | "Genetic and genomic tools for Xenopus research: The NIH Xenopus |
| RT | initiative;" |
| RV | Dev. Dyn. 225:384-391(2002). |
| RV | [3] |
| RV | SEQUENCE FROM N.A. |
| RP | TISSUE=Kidney; |
| RA | Klein S., Strausberg R., |
| RL | Submitted (JUN-2004) to the EMBL/genbank/DBD databases. |
| CC | -1 SIMILARITY: Belongs to the Ser/Thr protein kinase family. |
| DR | EMBL; BC074183; AAH74183.1; . |
| DR | InterPro; IPR011009; Kinase like. |
| DR | InterPro; IPR002290; Ser_Thr_kinase. |
| DR | InterPro; IPR008271; Ser_thr_pkin_AS. |
| DR | InterPro; IPR001245; Tyr_pkinase. |
| DR | Pfam; PF00069; Pkinase; 1. |
| DR | ProDom; PD000001; Prot_Kinase; 1. |
| DR | SMART; SMO0220; S_TKC; 1. |
| DR | SMART; SMO0219; TYKIC; 1. |
| DR | PROSITE; PS00107; PROTEIN_KINASE_ATP; 1. |
| DR | PROSITE; PS50011; PROTEIN_KINASE_DOM; 1. |
| DR | PROSITE; PS00108; PROTEIN_KINASE_ST; 1. |
| KW | ATP-binding; Hypothetical protein; Kinase; |
| KM | Serine/threonine-protein kinase; Transferase. |
| SO | SEQUENCE 395 aa; 43896 MW; 4C78BBBF828A2CC4 CRC64; |
| Query Match | 10.9%; Score 240.5; DB 2; Length 395; |
| Best Local Similarity | 27.4%; Pred. No. 7.5e-10; |
| Matches | .90; Conservative 48; Mismatches 108; Indels 83; Gaps 16 |
| OY | 4 SCGAMGAEPPEFLAEFGRLMVQSRLSGSSASV-----YYRCCSNPESPCQA 52 |
| Db | 9 SGACMKKK----VEDIKRMFOFKETGLTGAFSEVVLAEEKYTKLFAYVKI-----PKKA 59 |
| OY | 53 LRQLFPPPTTGAASAAYEGPFKEKPAALAEOLCGHNITLYGVFTIHSPNPSRCLL-- 111 |
| Db | 60 LK-----GKSSSIENEIAVLARKIK-HENIVALEDIY-----ESPSHLYLM 99 |
| OY | 112 -----ELLDVSVSELILYSSHQCMMMTIQHCADVYLEALAFLEHGEGVHADLPKNIL 165 |
| Db | 100 QLVSGGEFLFD-RIVEKGFYTEKDASTL-----IQOVDAYSYLRLGIYVRHDKPENLL 152 |
| OY | 166 WSAENECEKLI-I-DFGISFKEGNOVDKYIQ--TDGYRAPAEALONCLAQAGIQSDPECTS 221 |
| Db | 153 YFSQVEEKSIMIDSLGSKMEKGDMVSTACGTGGYVAPEV-----LAAQ-----KPYSK 201 |
| OY | 222 AVDLMSLGIIILEMWSGMKLTKTIVASQEKMANSSAIIDHIPASKAVVNAAIPAY-----H 276 |

| Db | 202 | ADWCMSIGVIAVILLGCG-----PPFYDENSRLFEQLL--KAQYFDSFYMWDISES | 252 |
|--|---|---|-----|
| Qy | 277 | LDLTKSMHDDPSRRIPAEWALCSPPSS 305 | |
| Db | 253 | ADDFIQNLMEKDPNRRRYTCEQALRHPWIA 281 | |
| RESULT 11. | | | |
| 09CA22 | 09CA22 | PRELIMINARY; PRT; 720 AA. | |
| AC | 09CA22; | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Created) | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last sequence update) | | |
| DT | 01-MAR-2004 (TREMBLrel. 26, Last annotation update) | | |
| DE | Hypothetical protein T32E8.5. | | |
| GN | Name=132E8.5; | | |
| OS | Arabidopsis thaliana (Mouse-ear cress). | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| OC | Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; rosids; | | |
| OC | eurosid II; Brassicales; Brassicaceae; Arabidopsis. | | |
| OX | NCBI_TaxID=3702; | | |
| RN | [1] | | |
| RA | SEQUENCE FROM N.A. | | |
| RA | Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D., | | |
| RA | Maiti R., Rongning C.M., Koo H., Fujii C.Y., Uteback T.R., | | |
| RA | Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.; | | |
| RL | Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases. | | |
| RN | [2] | | |
| RA | SEQUENCE FROM N.A. | | |
| RA | Town C.D., Kaul S.; | | |
| RL | Submitted (Jan-2001) to the EMBL/GenBank/DBJ databases. | | |
| DR | EMBL: AC012193; MAG51619.1; -. | | |
| DR | PIR; A96807; A96807. | | |
| DR | GO; GO:0005524; P:ATP binding; IEA. | | |
| DR | GO; GO:0004674; P:protein serine/threonine kinase activity; IEA. | | |
| DR | GO; GO:0016740; P:transferase activity; IEA. | | |
| DR | GO; GO:0006668; P:protein amino acid phosphorylation; IEA. | | |
| DR | InterPro; IPR011009; Kinase like. | | |
| DR | InterPro; IPR000719; Prot. Kinase. | | |
| DR | InterPro; IPR002290; Ser thr kinase. | | |
| DR | InterPro; IPR008271; Ser thr_pkin_AS. | | |
| DR | Pfam; PF00069; Pkinase; 1. | | |
| DR | Prodom; PD000001; Prot. kinase; 1. | | |
| DR | SMART; SM00220; S_TKc_1. | | |
| DR | PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1. | | |
| DR | PROSITE; PS00011; PROTEIN KINASE DOM; 1. | | |
| DR | PROSITE; PS01008; PROTEIN KINASE ST; UNKNOWN_1. | | |
| KM | APR-binding; Hypothetical protein; Kinase; Transferase. | | |
| SO | SEQUENCE 720 AA; 80027 MW; P26A9A0B1C6D91DE CRC64; | | |
| Query Match 10.9%; Score 240.5; DB 2; Length 720; | | | |
| Best local Similarity 25.7%; Pred. No. 1.6e-09; | | | |
| Matches 105; Conservative 57; Mismatch 146; Indels 101; Gaps 17; | | | |
| Qy | 11 | EPFRFLFAGRLMVQSHLSSGSSASVYRRCGSPSPGALRQLPPTGGAASAE 70 | |
| Db | 331 | DDPFLFKNVNGKLYQRLKIGSGSSSEVHKV--SSDCTIYALKKIKLGRDYAT- 383 | |
| Qy | 71 | YGRKERALAEQLQGRHRIVTLVGVFTIHSPVPSKCLLELDIVSE----- 120 | |
| Db | 384 | YGRCEIGELKTKLKKNTMIID-----IDYETLDTKLLQEVLTNGTMSKGRVKEGF 436 | |
| Qy | 121 | ---LLLYS---SHQGSQW-----MICHQARDVLEALFLHREGYHADIK 160 | |
| Db | 437 | IYVWLEYGIDLAHMLSQKMEIEGSDRTIDEMWLRPYWQQLIQAVNTIHEEIVHSDIK 496 | |
| Qy | 161 | PNRIILSAENEFKLIIDGSLFKSGNDVKKIQD-----GYRAEAEIQNLQAGIQ 214 | |
| Db | 497 | PANFL--LVKGLFKLIDGIA-KAINDDTNINQSDSQVGLTSLYSSEAPMCNSEDNG-- 551 | |
| Qy | 215 | SDTECSANVLMSGLIILLEMFSQMKLKHVRSQEWKANSATIID--HIFASKAVYNAAI 272 | |
| Db | 552 | NITKCGRPBDILSLSCILLYQVNYG-KRPAPADYKTF-AKKVITIDPNHETITNQSLNP-- 607 | |


```

QY 273 PAYHLRLIKSMHDDPSRRIPAEMLCSPPFSIPAPH----- 311
DB 608 ----WLDLWKCKCLAWBRNGWRIRPELLOHFLAPL-PHEPOVKITKLPSLIASGSGD 663
QY 312 -----IEDVLMPTPVLLRLNLVDDYLGNEEBEYEDVEDVEEC 351
DB 664 DKANSWISOLEQLSNPAPLPR-----NDVLDSRDQNOQLSRVSEL 706

RESULT 12
ID Q25758 PRELIMINARY; PRT; 765 AA.
AC Q25758;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE Mitogen-activated protein kinase-related protein.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dd2;
RA Lin D.T., Goldman N.D., Syin C.;
RT "Stage-specific expression of a Plasmodium falciparum protein related to the eukaryotic mitogen-activated protein kinases.";
RL Mol. Biochem. Parasitol. 78:0-0(10).
(2)
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dd2;
RA Lin D.T., Goldman N.D., Syin C.;
RT Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY. Belongs to the Ser/Thr protein kinase family.
DR EMBL: U36377; AAC47170.1; -.
DR HSSP: Q16539; IKV1.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006688; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR003527; MAP_kin.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS01351; MAPK; UNKOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 765 AA; 90231 MW; 5BEC8EA6A4B0232 CRC64;

Query Match 10.9%; Score 240.5; DB 2; Length 765;
Best Local Similarity 25.1%; Pred. No. 1.7e-09;
Matches 103; Conservative 57; Mismatches 145; Indels 105; Gaps 19;

QY 23 WQVQSLAGSSGSASVYVRCCGPGSP-----GALROFLPPTGTGAASAEVGRKRR 77
DB 23 YDLKKVKGAGVGVKGRCKKNKIVAVKKIGARQN-----CTDAQRTFR-EL 71
QY 78 AALBQLQGRNIVTLTGVFTIHSPVPSRCLELLLDVSVSL-----LLYSHQCSMW 133
DB 72 IFLYELNGHNTIKLMDVIR-----AKQDNNDIYLLFDMEETDLHEVIRADLLLEIHK----- 123
QY 134 MIOHCARDVLEALFLHHEGVADLKPRVILWSAENEC-FKLIDRGLS-----FKEGN 186
DB 124 --KTIIVQLIRAKTIHSGGLRDIKPSNLT--VNSECHIKVADGGLARISITHVNEK 179
QY 187 ODV-KYIQTGYRABEALQNCIAQAGSDTECTSAVDLWSLGIILLMEPSGMKL----- 241

```

```

DB 180 VPILTDVATRWYRAPEI-----LLGSTHYTEDVDWMSLGCINGELCGKPLFTG 229
QY 242 KHTVRSQE-----WKANSATIID--HIFASAAVNV-AAIPAYHLR-----DLI 281
DB 230 NSTMNOLEKIIVOTGPKNKDIEDIRSPAEKITISSPDLDKKKAKDICYRASNSLDLL 289
QY 282 KSMHDDPSRRIPAEML-----CSPFFSIPAPHIEDVLMPTPVLLRLNLVDDYLGNE 337
DB 290 EKLQFNPSKRIASNALMKHKKVVEEFGHSIIDPFCRHIITIP----- 331

QY 338 EYEDVEDVEECCQKCPGVSVSLVPEKNGRGQVFEYANAGDSKAAQK 387
DB 332 -----INDTKYRVNFYRNVYFVIMRRNKFHSNVL-----NQGSKKKEK 372

RESULT 13
ID Q84VX4 PRELIMINARY; PRT; 777 AA.
AC Q84VX4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, last annotation update)
DE Atlg77720.
OC Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eutroside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamlaya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Soultwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BR004636; AA042882.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006688; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKOWN_1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKOWN_1.
KW ATP-binding; Kinase; Transferase.
KW SEQUENCE 777 AA; 86376 MW; F2C25DA9609A878F CRC64;

Query Match 10.9%; Score 240.5; DB 2; Length 777;
Best Local Similarity 25.7%; Pred. No. 1.7e-09;
Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;

QY 11 EPPRFLEAFGRMVOQSRIGSSGSASVYVRCCGPGSPGALROFLPPTGTGAASAAE 70
DB 388 DPDLFFKNGKIKYQKLGKIGSGSSSEVKKVT---SSDCIYLYKKIKYKGRVYAR- 440
QY 71 YGFRKRAALEQLOQRNIVTLTGVFTIHSPVPSRCLELLLDVSVSE----- 120
DB 441 YGFCOEIGYKLLKKKTNITQL-----IDYEVTDKTLDEVLNGTWSNKKDGRVKEGDF 493
QY 121 ---LLYS-----SHQCSMW-----MIOHCARDVLEALFLHHEGVYHADLK 160
DB 494 IYVLELYGSDILAHMLSQKWRBEIGSDRTIDENMWLFYWOQLQAVNTHERRIVHSDK 553
QY 161 PNIILWSAENECFKLIDRGLSFKEGNQDVKYIQT-----GYRABEALQNCIAQAGIQ 214

```

```

Db      554 PANFL--LVKGFPLKLDIFGIA--KAINSPTTNIDRDSQVGLTSYMSPEAFWCNESDENG-- 608
QY      215 SDRECSAVDLMSIGIILEMFSQMKLKHIVRSQEWKANSATIID--HIFASAVVNAAI 272
Db      609 NPTKCRPSISLGLGLTYQVYTG-RTPPADYKTFW-AKFKVITDPNHETITNQSLNP-- 664
QY      273 PAVHLRDLIKSMHDDPSRRIRPAEMALCSPPFSIPFAPH----- 311
Db      665 ---WLDLWKCKCLAMPBQNRQWRIRPELQHPFLAPL-PHEPQVKTIKIFSLIASEGSD 720
QY      312 -----IEDLVMLPTPYLLNLVDDDYLCNESEYEDVEDVKEEC 351
Db      721 DKANSMISQLEQLNSPAPLPR-----NDVIDRSDQNOQLSRVSELC 763

RESULT 14
ID      094656 PRELIMINARY; PRT; 826 AA.
AC      094656;
DT      01-FEB-1997 (TEMBLrel. 02, Created)
DT      01-FEB-1997 (TEMBLrel. 02, Last sequence update)
DE      01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE      Mitogen-activated protein kinase 1, serine/threonine protein
       kinase.
GN      Name=pfmap1;
OS      Plasmodium falciparum.
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=5833;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=IT04;
RA      Doering C.;
RL      Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR      EMBL; X82646; CNA57972.1; -.
DR      PIR; JCS153; JCS153.
DR      HSSP; Q16539; 1KV1.
DR      GO; GO:000524; F:ATP binding; IEA.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR003527; MAP kin.
DR      InterPro; IPR000719; Prot_kinase.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      InterPro; IPR008271; Ser_thr_pkin_AS.
DR      Pfam; PF00069; Pkinase; 1.
DR      SMART; SM00220; S_TKc; 1.
DR      PROSITE; PS01351; MAPK; UNKNOWN 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR      ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW      SEQUENCE 826 AA; 96645 MW; A68BEF7E734A38A CRC64;

Query Match      10.98; Score 240.5; DB 2; Length 826;
Best Local Similarity 25.14; Pred. No. 1.9e-09;
Matches 103; Conservative 57; Mismatches 145; Indels 105; Gaps 19;
QY      23 WQVSRILSGSSASVYRVRCGNGPSPP-----GALRQFLPPTGTGAASAAYGFRKER 77
Db      23 YDLIKVVGKAGVGVFVRGRCKKKNIVAVKIRGATQN-----CTDAQRTFR-EL 71

```

```

QY      78 AALBQLQGRNIVLYGVFTTHSPNVSRCLLELLVSVSEL-----LIVSHQGCMM 133
Db      72 IFLYEINGDNIITKIMDVYK--AKNDNDIYLIPFMEITDHEVIKADLEIRK----- 123
QY      134 MIOHCARDVLEALFLHHEGVHADLKPRNITMSAENEC-FGLIDFGS-----FKEGN 186
Db      124 --KVIITVQLRALKTIHSGGLIRDKFSNII--VNSECHIKVADFGIARISITHVENK 179
QY      187 QDV--KYIOTDGYRAPEALQNCIAQIQSPTCEASVDLMSIGIILEMFSQMKL--- 241
Db      180 VPILDYVATRWYRAPEI-----LIGSTHYTEDVDMWSIGCINGELLGKPLFTG 229
QY      242 KATVRSQ-----WKANSATIID--HIFASAVVN-AAIPAYHNR-----DLI 281
Db      230 NSTMNQLKTIQVICKPKPKDIEDIRSPFAEKTISSFVDLKKKDKIDCYRASNSLDL 289
QY      282 KSMHDDPSRRIRPAEMAL---CSPFSPFAPHIEDLVMLPTPYLLNLVDDDYLCNE 337
Db      290 EKLQFPSPKRLSAENALNKHKVEEFSHIIDEPFRHIIITP----- 331
QY      338 EEYEDVEDVKEECQKYGPVVSLVPEKPNRGQVFEVYANAGSKAAQK 387
Db      332 -----INDTKYRVNPFYRNVVYFVIMRRNKKFHSNVL-----NGSEAKKEK 372

RESULT 15
ID      081LF0 PRELIMINARY; PRT; 914 AA.
AC      081LF0;
DT      01-MAR-2003 (TEMBLrel. 23, Created)
DT      01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE      01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE      Mitogen-activated protein kinase 1.
GN      ORFNames=PF14_0294;
OS      Plasmodium falciparum (isolate 3D7).
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=36329;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=22255705; PubMed=12368864;
RA      Gardner M.J., Hall N., Fung E., White O., Bertlman M., Hyman R.W.,
RA      Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA      Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kye S.,
RA      Chan M.S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA      Perlea M., Allen J., Selengut J., Haft J., Mather M.W., Vaidya A.B.,
RA      Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA      McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA      Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA      Frazer C.M., Barrall B.;
RT      "Genome sequence of the human malaria parasite Plasmodium
       falciparum."
RL      Nature 419:498-511 (2002).
CC      -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR      EMBL; AE014820; AAN36907.1; -.
DR      HSSP; Q16539; 1KV1.
DR      GO; GO:000524; F:ATP binding; IEA.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR003527; MAP kin.
DR      InterPro; IPR000719; Prot_kinase.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      InterPro; IPR008271; Ser_thr_pkin_AS.
DR      Pfam; PF00069; Pkinase; 1.
DR      SMART; SM00220; S_TKc; 1.
DR      PROSITE; PS01351; MAPK; UNKNOWN 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR      ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW      SEQUENCE 914 AA; 107277 MW; CA4F486AB051F472 CRC64;

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2004, 11:26:47 ; Search time 109.5 Seconds
(without alignments)
1372.672 Million cell updates/sec

Title: US-10-798-532-4

Perfect score: 2206

Sequence: 1 MAGSCGAMGAEPRFLFARF.....VATFYPLSAVKRGYLTLL 419

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 2195 | 99.5 | 419 | 4 | AAG65767 Human pro |
| 2 | 2195 | 99.5 | 419 | 5 | ABP64797 Human pro |
| 3 | 2190.5 | 99.3 | 418 | 3 | AAV70306 Human tra |
| 4 | 2188 | 99.2 | 419 | 4 | AAU03534 Human pro |
| 5 | 2188 | 99.2 | 419 | 4 | AAE11769 Human kin |
| 6 | 2187.5 | 99.2 | 418 | 3 | AAV70305 Human wil |
| 7 | 2183 | 99.0 | 419 | 7 | ADD48093 Rat Prote |
| 8 | 483.5 | 21.9 | 137 | 5 | ABP64662 Human ORF |
| 9 | 240.5 | 10.9 | 638 | 3 | AAG38502 Arabidops |
| 10 | 240.5 | 10.9 | 720 | 3 | AAG38501 Arabidops |
| 11 | 240.5 | 10.9 | 777 | 3 | AAG38500 Arabidops |
| 12 | 236.5 | 10.7 | 544 | 2 | AAW85028 CDR2-gree |
| 13 | 235 | 10.7 | 298 | 7 | ADD21398 Golden ha |
| 14 | 234.5 | 10.6 | 433 | 3 | AAH19094 Protein e |
| 15 | 234.5 | 10.6 | 433 | 8 | ADP87472 S cerevis |
| 16 | 234 | 10.6 | 376 | 3 | AAAG04672 Arabidops |
| 17 | 233 | 10.6 | 435 | 6 | AAE31916 Yeast Hog |
| 18 | 232.5 | 10.5 | 367 | 3 | AAAB2137 Rat stres |
| 19 | 232 | 10.5 | 435 | 6 | AAE31917 Yeast Hog |
| 20 | 232 | 10.5 | 435 | 6 | AAE31922 Yeast Hog |
| 21 | 232 | 10.5 | 435 | 6 | AAE31921 Yeast Hog |
| 22 | 232 | 10.5 | 435 | 6 | AAE31914 Yeast Hog |
| 23 | 232 | 10.5 | 435 | 6 | AAE31913 Yeast Hog |
| 24 | 232 | 10.5 | 544 | 2 | AAW85029 Green flu |
| 25 | 232 | 10.5 | 568 | 2 | AAW75791 Human pro |

| | | | | | |
|----|-------|------|-----|---|---------------------|
| 26 | 232 | 10.5 | 568 | 3 | AAAB0059 hYAK3-2 p |
| 27 | 232 | 10.5 | 568 | 4 | AAE02013 Human YAK |
| 28 | 232 | 10.5 | 587 | 7 | AAH57639 Human DYR |
| 29 | 232 | 10.5 | 588 | 2 | AAW75790 Human pro |
| 30 | 232 | 10.5 | 588 | 3 | AAH68792 Amino aci |
| 31 | 232 | 10.5 | 588 | 3 | AAAB0058 hYAK3-2 a |
| 32 | 232 | 10.5 | 588 | 4 | AAE02012 Human YAK |
| 33 | 232 | 10.5 | 801 | 7 | ADD93285 p53-Yak3 |
| 34 | 231.5 | 10.5 | 371 | 3 | AAAG45420 Arabidops |
| 35 | 231.5 | 10.5 | 384 | 2 | AAH89411 Stress ac |
| 36 | 231.5 | 10.5 | 426 | 2 | AAH89410 Stress ac |
| 37 | 231.5 | 10.5 | 426 | 5 | ABH85009 Pain regu |
| 38 | 231.5 | 10.5 | 549 | 8 | AD145327 Rice isop |
| 39 | 231 | 10.5 | 298 | 2 | AAW56689 Homo sap1 |
| 40 | 231 | 10.5 | 298 | 7 | ADD21402 Chinese h |
| 41 | 231 | 10.5 | 298 | 7 | ADD21406 Rat cycl1 |
| 42 | 231 | 10.5 | 298 | 7 | ADD21394 Mouse cyc |
| 43 | 231 | 10.5 | 416 | 2 | AAW06795 High Osmo |
| 44 | 231 | 10.5 | 435 | 6 | AAE31910 Yeast Hog |
| 45 | 231 | 10.5 | 435 | 6 | AAE31915 Yeast Hog |

ALIGNMENTS

RESULT 1
ID AAG65767 standard; protein; 419 AA.
XX
AC AAG65767;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human protein kinase polypeptide 13887.
XX
KW Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human;
KM cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatotrophic;
KM gene therapy.
XX
OS Homo sapiens.
XX
PN WC0200173050-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001WC-US009483.
XX
PR 24-MAR-2000; 2000US-0191846P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers R;
XX
DR WPI; 2001-611632/70.
DR N-PSDB; AAI66828, AAI66829.
XX
PT New human protein kinase polypeptides; 3714, 16742, 23546 and 13887,
PT useful in diagnosis of cancer or cellular proliferation or
PT differentiation disorders and to screen for polypeptide modulators useful
PT to treat such conditions.
XX
PS Claim 4; Fig 10A-C; 169pp; English.
XX
CC The invention provides novel human protein kinase polypeptides; 3714,
CC 16742, 23546 and 13887 and nucleic acid molecules encoding them. The
CC protein kinase polypeptides can be expressed by standard recombinant
CC methodology; 3714, 16742, 23546 or 13887 nucleic acids and polypeptides
CC are useful for diagnostic and screening methods to identify subjects (at
CC risk of) having cancer or cellular proliferation and/or differentiation
CC disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and
CC modulators are useful for the treatment of cancer, particularly colon
CC cancer or cellular proliferation and/or differentiation disorders. Other
CC disorders associated with 3714, 16742, 23546 or 13887 expression or

CC activity that can be treated include bone related disorders, inflammatory
 CC disorders, autoimmune diseases, cardiovascular disorders and liver
 CC diseases. The present sequence represents a human protein kinase
 CC polypeptide 13867

XX Sequence 419 AA;

Query Match 99.5%; Score 2195; DB 4; Length 419;
 Best Local Similarity 99.5%; Pred. No. 5.3e-229;
 Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPPRFLFAFGRLMVOVSRLGSSSASVYRRCGNGSPGALRQFLPPG 60
 DB 1 MAGSGCAGWAEPPRFLFAFGRLMVOVSRLGSSSASVYRRCGNGSPGALRQFLPPG 60
 QY 61 TTGAASAAEYGRKERALBQLOGHNTVLTLYGVFTIHSPVPSRCLLELDVSVSE 120
 DB 61 TTGAASAAEYGRKERALBQLOGHNTVLTLYGVFTIHSPVPSRCLLELDVSVSE 120
 QY 121 LLLYSSHQSCMMWIOHCARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLIDFGL 180
 DB 121 LLLYSSHQSCMMWIOHCARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLIDFGL 180
 QY 122 LLLYSSHQSCMMWIOHCARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLIDFGL 180
 DB 122 LLLYSSHQSCMMWIOHCARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLIDFGL 180
 QY 181 SFKEGQDVYKTIOTDGRAPAEALONCLAQAGLOSTECTSAVDLWSLGIILLEMPSGK 240
 DB 181 SFKEGQDVYKTIOTDGRAPAEALONCLAQAGLOSTECTSAVDLWSLGIILLEMPSGK 240
 QY 241 LKHTVSOEWMKNSAIIDHIFASKAVVNAIPAYHLRDLIKSMLHDDPSRRIPEMALC 300
 DB 241 LKHTVSOEWMKNSAIIDHIFASKAVVNAIPAYHLRDLIKSMLHDDPSRRIPEMALC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDYLGNEEYEDVDEVEKCEQKYGVSVL 360
 DB 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDYLGNEEYEDVDEVEKCEQKYGVSVL 360
 QY 361 LVPEKPNRGQVFEVYANAGDSKAAQKLLTGMPDGFVATFYPLISAYKRGYLTQL 419
 DB 361 LVPEKPNRGQVFEVYANAGDSKAAQKLLTGMPDGFVATFYPLISAYKRGYLTQL 419

RESULT 2
 ABP64797
 ID ABP64797 standard; protein; 419 AA.

XX ABP64797;
 AC
 XX
 DT 25-FEB-2003 (first entry)
 DE
 XX Human protein SEQ ID 457.

XX Human; expressed sequence tag; EST; haematopoietic disorder;
 KW central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; noctropic;
 KW anti-allergic; anti-inflammatory; immunosuppressive; neutropenic;
 KW cytotoxic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.

XX Homo sapiens.
 OS
 XX
 PN WO200259260-A2.

PD 01-AUG-2002.

XX 16-NOV-2001; 2001MO-US042950.

PR 17-NOV-2000; 2000US-00714936.

XX (HYSE-) HYSBQ INC.

XX Tang YT, Goodrich RM, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Yang Y, Wehman T, Drmanac RT;

XX WPI; 2002-590824/63.
 DR N-PSDB; ABQ99383.
 XX
 PT New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.

Claim 20; SEQ ID NO 457; 394bp; English.

CC The present invention relates to novel human coding sequences (ABQ99268-
 CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic, and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotent or pluripotent state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPRO at
 CC ftp.wipro.int/pub/published_pct_sequences

XX Sequence 419 AA;

Query Match 99.5%; Score 2195; DB 5; Length 419;
 Best Local Similarity 99.5%; Pred. No. 5.3e-229;
 Matches 417; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGSGCAGWAEPPRFLFAFGRLMVOVSRLGSSSASVYRRCGNGSPGALRQFLPPG 60
 DB 1 MAGSGCAGWAEPPRFLFAFGRLMVOVSRLGSSSASVYRRCGNGSPGALRQFLPPG 60
 QY 61 TTGAASAAEYGRKERALBQLOGHNTVLTLYGVFTIHSPVPSRCLLELDVSVSE 120
 DB 61 TTGAASAAEYGRKERALBQLOGHNTVLTLYGVFTIHSPVPSRCLLELDVSVSE 120
 QY 121 LLLYSSHQSCMMWIOHCARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLIDFGL 180
 DB 121 LLLYSSHQSCMMWIOHCARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLIDFGL 180
 QY 122 LLLYSSHQSCMMWIOHCARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLIDFGL 180
 DB 122 LLLYSSHQSCMMWIOHCARDVLEALFLHHEGYVHADLKPNNILMSAENECFKLIDFGL 180
 QY 181 SFKEGQDVYKTIOTDGRAPAEALONCLAQAGLOSTECTSAVDLWSLGIILLEMPSGK 240
 DB 181 SFKEGQDVYKTIOTDGRAPAEALONCLAQAGLOSTECTSAVDLWSLGIILLEMPSGK 240
 QY 241 LKHTVSOEWMKNSAIIDHIFASKAVVNAIPAYHLRDLIKSMLHDDPSRRIPEMALC 300
 DB 241 LKHTVSOEWMKNSAIIDHIFASKAVVNAIPAYHLRDLIKSMLHDDPSRRIPEMALC 300
 QY 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDYLGNEEYEDVDEVEKCEQKYGVSVL 360
 DB 301 SPFFSIPFAPHIEDLVMLPTPVRLNLNVDDYLGNEEYEDVDEVEKCEQKYGVSVL 360
 QY 361 LVPEKPNRGQVFEVYANAGDSKAAQKLLTGMPDGFVATFYPLISAYKRGYLTQL 419
 DB 361 LVPEKPNRGQVFEVYANAGDSKAAQKLLTGMPDGFVATFYPLISAYKRGYLTQL 419

RESULT 3

```

AAU0306
ID   AAY70306 standard; protein; 418 AA.
AC   AAY70306;
XX
XX
DT   06-JUN-2000 (first entry)
XX
XX   Human transdominant mutant serine/threonine kinase KIS (hKIS) protein.
XX
XX   KIS; human; hKIS; serine/threonine kinase; transdominant mutant; p27;
XX   G1 phase; cell proliferation; cyclin-dependent kinase inhibitor; CKI;
XX   modulator; treatment; cell proliferative disease; vascular disorder;
XX   gene therapy; atherosclerosis; restenosis.
XX
XX   Homo sapiens.
XX
XX   Key Location/Qualifiers
XX   FT   Misc-difference 54 /note= "Wild type Lys substituted with Arg"
XX   FT   Misc-difference 186 /note= "Encoded by AATCAG"
XX   FT   Misc-difference 186 /note= "Encoded by AATCAG"
XX
XX   WO200011165-A1.
XX
XX   02-MAR-2000.
XX
XX   20-AUG-1999; 99WO-US018903.
XX
XX   21-AUG-1998; 98US-0097710P.
XX
XX   (NABE/) NABEL G J.
XX   (NABE/) NABEL E G.
XX
XX   Nabel GJ, Nabel EG;
XX
XX   WPI; 2000-237648/20.
XX   DR   N-PSDB; AAZ51356.
XX
XX   Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
XX   PT   for inhibiting the cyclin kinase inhibitor p27, and so alter cell
XX   PT   proliferation.
XX
XX   Example 1; Page 59-61; 70pp; English.
XX
XX   The present protein sequence is the transdominant mutant human KIS
XX   CC   (hKIS), constructed by site directed mutagenesis. A single nucleotide
XX   CC   substitution (A to G) results in a lysine to arginine change in the
XX   CC   protein sequence. hKIS is a serine/threonine kinase, that acts as an
XX   CC   inhibitory kinase of cyclin-dependent kinase inhibitor (CKI), p27. hKIS
XX   CC   controls cell proliferation and is localised predominantly in the
XX   CC   nucleus. It binds to CKI p27 and inhibits its ability to arrest cells in
XX   CC   G1 phase. The hKIS sequences are used to modulate cell proliferation and
XX   CC   treat cell proliferative and vascular diseases. The polynucleotide
XX   CC   sequence may be used in gene therapy to treat vascular disorders such as
XX   CC   restenosis or atherosclerosis
XX
XX   Sequence 418 AA;
XX
XX   Query Match 99.3%; Score 2190.5; DB 3; Length 418;
XX   Best Local Similarity 99.8%; Pred. No. 1.6e-228;
XX   Matches 418; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
XX
XX   1 MAGSGCMAEPRPRLEAFGRMLQVQSRGLGSSASAYRRCCGNGSPPGALRQFLPPG 60
XX   1 MASSGGAMGAEPRPRLEAFGRMLQVQSRGLGSSASAYRRCCGNGSPPGALRQFLPPG 60
XX
XX   61 TTGAASAAEYGRKERRALEQOGHRNITLKGVTTHSPVNPVSRCLLELLDVSVSE 120
XX   61 TTGAASAAEYGRKERRALEQOGHRNITLKGVTTHSPVNPVSRCLLELLDVSVSE 120
XX   61 TTGAASAAEYGRKERRALEQOGHRNITLKGVTTHSPVNPVSRCLLELLDVSVSE 120
XX   61 TTGAASAAEYGRKERRALEQOGHRNITLKGVTTHSPVNPVSRCLLELLDVSVSE 120
XX
XX   121 LLIISSHQGSSMMIOHCARDVLEALAFIHGEGVHADLKPRNIIWSAENECEKFLIDFGL 180
XX   121 LLIISSHQGSSMMIOHCARDVLEALAFIHGEGVHADLKPRNIIWSAENECEKFLIDFGL 180
XX   121 LLIISSHQGSSMMIOHCARDVLEALAFIHGEGVHADLKPRNIIWSAENECEKFLIDFGL 180

```

```

XX   181 SPREGNOVYKIQTGDYRAPEALONCLAQAGLSQDTECTSAVDLSIGIILLEMFGMK 240
XX   181 SPREGN-DVKYIQTGDYRAPEALONCLAQAGLSQDTECTSAVDLSIGIILLEMFGMK 239
XX
XX   241 LKHTVRSQEWKANSNAIIDHIFASKAVVNAIIPAYHLRDLIKSLHDDPSRRIPAEMLC 300
XX   240 LKHTVRSQEWKANSNAIIDHIFASKAVVNAIIPAYHLRDLIKSLHDDPSRRIPAEMLC 299
XX
XX   301 SPFFSIPFAPHIEDLVMLPTPLRLNLVDDDYLGNEEBEYEDVDEECCQKGPVSL 360
XX   300 SPFFSIPFAPHIEDLVMLPTPLRLNLVDDDYLGNEEBEYEDVDEECCQKGPVSL 359
XX
XX   361 LVPKENPGRGQVFVEYANAGDSKAOKLLTGRMPDGKFPVATFPLSAKRGVLYQTLL 419
XX   360 LVPKENPGRGQVFVEYANAGDSKAOKLLTGMFPGKFPVATFPLSAKRGVLYQTLL 418
XX
XX   RESULT 4
XX   AAU03534
XX   ID   AAU03534 standard; protein; 419 AA.
XX
XX   AAU03534;
XX
XX   12-SEP-2001 (first entry)
XX
XX   Human protein kinase #34.
XX
XX   Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
XX   XX   metabolic disorder; immune related disease; neurological disorder;
XX   XX   neurodegenerative disorder; inflammatory disorder; infectious disease;
XX   XX   reproductive disorder.
XX
XX   Homo sapiens.
XX
XX   WO200138503-A2.
XX
XX   31-MAY-2001.
XX
XX   22-NOV-2000; 2000WO-US032085.
XX
XX   24-NOV-1999; 99US-0167482P.
XX
XX   (SUGR-) SUGEN INC.
XX
XX   Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX   PI   Flanagan P, Clary D;
XX
XX   WPI; 2001-343950/36.
XX   DR   N-PSDB; AAS06734.
XX
XX   Nucleic acids encoding human kinase polypeptides, useful for preventing
XX   PT   diagnosing and/or treating e.g. cancer, immune, cardiovascular and
XX   PT   neuronal-associated diseases, and microbial infections.
XX
XX   Claim 7; Fig 2; 43pp; English.
XX
XX   AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
XX   CC   protein kinases have been identified as members of the tyrosine or
XX   CC   serine/threonine kinase (PTK and STK) families. The polynucleotides
XX   CC   encoding protein kinases and the polypeptides may be used in the
XX   CC   prevention, diagnosis and treatment of diseases associated with
XX   CC   inappropriate kinase expression. For example, they may be used to treat
XX   CC   cancers (especially cancers of haematopoietic origin), cardiovascular
XX   CC   disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
XX   CC   immune related diseases (e.g. rheumatoid arthritis), neurological
XX   CC   disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
XX   CC   Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
XX   CC   disease (e.g. HIV) and reproductive disorders (e.g. infertility).
XX   CC   Additionally, polynucleotides encoding protein kinases may be used for
XX   CC   gene therapy and as DNA probes in diagnostic assays. The protein kinase
XX   CC   polypeptides may be used as antigens in the production of antibodies
XX   CC   against the protein kinases and in assays to identify modulators of

```

[illegible]

| Query Match | Score | DB 4 | Length | 419 |
|-----------------------|---|--------------------|----------|--------|
| Best Local Similarity | 99.3% | Pred. No. 3.1e-228 | | |
| Matches 416 | Conservative 1 | Mismatches 2 | Indels 0 | Gaps 0 |
| 1 | MAGSGCAAGAPPPFLLEAFGLTMOVQSLGSSSSAYVRVCCGPGSPPGALROFLPPG | 60 | | |
| 1 | | | | |
| 1 | MAGSGCAAGAPPPFLLEAFGLTMOVQSLGSSSSAYVRVCCGPGSPPGALROFLPPG | 60 | | |
| 1 | | | | |
| 1 | TTGAASAAEYGFPRKRAALFQLOGHRNIIVLYGVFTIHFSPPNVSRCLLELDIVSYE | 120 | | |


```

Db      61 TTGAASAAEYGRKERAALEQLOGHRNIVTLGVFTIHSPVPSRCLLELDVSVSE 120
Qy      121 LLLYSSHQCGSMWMIQHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Db      121 LLLYSSHQCGSMWMIQHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Qy      181 SFKEGNQDVYKIQTDGYRAPEAELOQCLAOAGLOSDTECTSAVDLMSLGIILLEMFSGMK 240
Db      181 SFKEGNQDVYKIQTDGYRAPEAELOQCLAOAGLOSDTECTSAVDLMSLGIILLEMFSGMK 240
Qy      241 LKHTVRSQEWKANSALIDHIFASKAVVNAIPAYHLRDLIKSMIHDHDPSSRIIPAEWALC 300
Db      241 LKHTVRSQEWKANSALIDHIFASKAVVNAIPAYHLRDLIKSMIHDHDPSSRIIPAEWALC 300
Qy      301 SPFFSIPFAPHIDVLMPTPVRLNLVDDYLGNEEYEDVVEDEKCECKYGPVSL 360
Db      301 SPFFSIPFAPHIDVLMPTPVRLNLVDDYLGNEEYEDVVEDEKCECKYGPVSL 360
Qy      361 LVPKPNRGQVFEVYANAGDSKAAOKLITGRMFDGKFFVATFPLSAIKRGYLYOTLL 419
Db      361 LVPKPNRGQVFEVYANAGDSKAAOKLITGRMFDGKFFVATFPLSAIKRGYLYOTLL 419

RESULT 6
AA70305
ID AA70305 standard; protein; 418 AA.
XX
AC AA70305;
XX
DT 06-JUN-2000 (first entry)
XX
DE Human wild type serine/threonine kinase KIS (hKIS) protein.
XX
KW KIS; human; hKIS; serine/threonine kinase; cell proliferation; G1 phase;
KW cyclin-dependent kinase inhibitor; CKI; p27; modulator; treatment;
KW cell proliferative disease; vascular disorder; gene therapy; restenosis;
KW atherosclerosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 186 /note="Encoded by AATCAG"
FT
XX
PN MO20001165-A1.
XX
PD 02-MAR-2000.
XX
PF 20-AUG-1999; 99WC-US018903.
XX
PR 21-AUG-1998; 98US-0097710P.
XX
PA (NABEL/) NABEL G J.
PA (NABEL/) NABEL E G.
XX
PI Nabel GJ, Nabel EG;
XX
DR WPI, 2000-237648/20.
DR N-PSDB; AAZ51355.
XX
PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides used
PT for inhibiting the cyclin kinase inhibitor p27, and so alter cell
PT proliferation.
XX
PS Example 1; Page 57-58; 70pp; English.
XX
CC The present protein sequence is the wild type human KIS (hKIS), isolated
CC from a human B-cell library, using a yeast two hybrid screening system.
CC hKIS is a serine/threonine kinase, that acts as an inhibitory kinase of
CC cyclin-dependent kinase inhibitor (CKI), p27. hKIS controls cell
CC proliferation and is localised predominantly in the nucleus. It binds to
CC CKI p27 and inhibits its ability to arrest cells in G1 phase. It has an

```

```

CC amino terminal serine/threonine kinase consensus region and a carboxy
CC terminus with 42% sequence similarity to huzAF5, a subunit of the
CC splicing factor U2AF. The hKIS sequences are used to modulate cell
CC proliferation and treat cell proliferative and vascular diseases. The DNA
CC sequence may be used in gene therapy to treat vascular disorders such as
CC restenosis or atherosclerosis
XX
SQ Sequence 418 AA;
Query Match 99.2%; Score 2187.5; DB 3; Length 418;
Best Local Similarity 99.5%; Pred. No. 3.5e-228;
Matches 417; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
Qy      1 MAGSGCAMEPRPFLFAGRLMWQVSRGSGSSAVYVRCNGNPGSPGALROPLPG 60
Db      1 MAGSGCAMEPRPFLFAGRLMWQVSRGSGSSAVYVRCNGNPGSPGALROPLPG 60
Qy      61 TTGAASAAEYGRKERAALEQLOGHRNIVTLGVFTIHSPVPSRCLLELDVSVSE 120
Db      61 TTGAASAAEYGRKERAALEQLOGHRNIVTLGVFTIHSPVPSRCLLELDVSVSE 120
Qy      121 LLLYSSHQCGSMWMIQHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Db      121 LLLYSSHQCGSMWMIQHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Qy      121 LLLYSSHQCGSMWMIQHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Db      121 LLLYSSHQCGSMWMIQHCARDVLEALAFHHEGYVADLKPRNIIWSAENECFKLIDFGL 180
Qy      181 SFKEGNQDVYKIQTDGYRAPEAELOQCLAOAGLOSDTECTSAVDLMSLGIILLEMFSGMK 240
Db      181 SFKEGN-DVKYIQTDTGYRAPEAELOQCLAOAGLOSDTECTSAVDLMSLGIILLEMFSGMK 239
Qy      241 LKHTVRSQEWKANSALIDHIFASKAVVNAIPAYHLRDLIKSMIHDHDPSSRIIPAEWALC 300
Db      240 LKHTVRSQEWKANSALIDHIFASKAVVNAIPAYHLRDLIKSMIHDHDPSSRIIPAEWALC 299
Qy      301 SPFFSIPFAPHIDVLMPTPVRLNLVDDYLGNEEYEDVVEDEKCECKYGPVSL 360
Db      300 SPFFSIPFAPHIDVLMPTPVRLNLVDDYLGNEEYEDVVEDEKCECKYGPVSL 359
Qy      361 LVPKPNRGQVFEVYANAGDSKAAOKLITGRMFDGKFFVATFPLSAIKRGYLYOTLL 419
Db      360 LVPKPNRGQVFEVYANAGDSKAAOKLITGRMFDGKFFVATFPLSAIKRGYLYOTLL 418

RESULT 7
ADD48093
ID ADD48093 standard; protein; 419 AA.
XX
AC ADD48093;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AAC53031, SEQ ID NO 13791.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN MO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002MO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333947P.
XX
PA (GENO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.

```


| | | | | | |
|----|----------------------|--|----|--------------|---------------|
| XX | AC | AG38502; | PR | 18-JUN-1999; | 99US-01394621 |
| XX | AC | | PR | 18-JUN-1999; | 99US-01394631 |
| XX | DT | 18-OCT-2000 (first entry) | PR | 18-JUN-1999; | 99US-01397501 |
| XX | XX | | PR | 18-JUN-1999; | 99US-01397631 |
| XX | XX | | PR | 21-JUN-1999; | 99US-01398171 |
| DE | Arabidopsis thaliana | protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence. | PR | 22-JUN-1999; | 99US-01398991 |
| XX | XX | | PR | 23-JUN-1999; | 99US-01403531 |
| XX | OS | Arabidopsis thaliana. | PR | 23-JUN-1999; | 99US-01403541 |
| XX | XX | | PR | 23-JUN-1999; | 99US-01403551 |
| XX | XX | | PR | 24-JUN-1999; | 99US-01406951 |
| XX | XX | | PR | 28-JUN-1999; | 99US-01408231 |
| XX | XX | | PR | 29-JUN-1999; | 99US-01409911 |
| XX | XX | | PR | 30-JUN-1999; | 99US-01412871 |
| XX | XX | | PR | 01-JUL-1999; | 99US-01418421 |
| XX | XX | | PR | 01-JUL-1999; | 99US-01421541 |
| XX | XX | | PR | 02-JUL-1999; | 99US-01420551 |
| XX | XX | | PR | 06-JUL-1999; | 99US-01423901 |
| XX | XX | | PR | 08-JUL-1999; | 99US-01428031 |
| XX | XX | | PR | 09-JUL-1999; | 99US-01429201 |
| XX | XX | | PR | 12-JUL-1999; | 99US-01429771 |
| XX | XX | | PR | 13-JUL-1999; | 99US-01435421 |
| XX | XX | | PR | 14-JUL-1999; | 99US-01436241 |
| XX | XX | | PR | 15-JUL-1999; | 99US-01440051 |
| XX | XX | | PR | 16-JUL-1999; | 99US-01440851 |
| XX | XX | | PR | 16-JUL-1999; | 99US-01440861 |
| XX | XX | | PR | 19-JUL-1999; | 99US-01443251 |
| XX | XX | | PR | 19-JUL-1999; | 99US-01443311 |
| XX | XX | | PR | 19-JUL-1999; | 99US-01443321 |
| XX | XX | | PR | 19-JUL-1999; | 99US-01443331 |
| XX | XX | | PR | 19-JUL-1999; | 99US-01443341 |
| XX | XX | | PR | 19-JUL-1999; | 99US-01443351 |
| XX | XX | | PR | 20-JUL-1999; | 99US-01443521 |
| XX | XX | | PR | 20-JUL-1999; | 99US-01443631 |
| XX | XX | | PR | 20-JUL-1999; | 99US-01446841 |
| XX | XX | | PR | 21-JUL-1999; | 99US-01448141 |
| XX | XX | | PR | 21-JUL-1999; | 99US-01450861 |
| XX | XX | | PR | 21-JUL-1999; | 99US-01450881 |
| XX | XX | | PR | 22-JUL-1999; | 99US-01450851 |
| XX | XX | | PR | 22-JUL-1999; | 99US-01450871 |
| XX | XX | | PR | 22-JUL-1999; | 99US-01450891 |
| XX | XX | | PR | 22-JUL-1999; | 99US-01451521 |
| XX | XX | | PR | 23-JUL-1999; | 99US-01451551 |
| XX | XX | | PR | 23-JUL-1999; | 99US-01452181 |
| XX | XX | | PR | 23-JUL-1999; | 99US-01452241 |
| XX | XX | | PR | 26-JUL-1999; | 99US-01452761 |
| XX | XX | | PR | 27-JUL-1999; | 99US-01455131 |
| XX | XX | | PR | 27-JUL-1999; | 99US-01455181 |
| XX | XX | | PR | 27-JUL-1999; | 99US-01455191 |
| XX | XX | | PR | 28-JUL-1999; | 99US-01455911 |
| XX | XX | | PR | 28-AUG-1999; | 99US-01463861 |
| XX | XX | | PR | 02-AUG-1999; | 99US-01463881 |
| XX | XX | | PR | 02-AUG-1999; | 99US-01463891 |
| XX | XX | | PR | 03-AUG-1999; | 99US-01470381 |
| XX | XX | | PR | 04-AUG-1999; | 99US-01472041 |
| XX | XX | | PR | 04-AUG-1999; | 99US-01473021 |
| XX | XX | | PR | 05-AUG-1999; | 99US-01471921 |
| XX | XX | | PR | 05-AUG-1999; | 99US-01472601 |
| XX | XX | | PR | 06-AUG-1999; | 99US-01473031 |
| XX | XX | | PR | 06-AUG-1999; | 99US-01474161 |
| XX | XX | | PR | 09-AUG-1999; | 99US-01474931 |
| XX | XX | | PR | 09-AUG-1999; | 99US-01475351 |
| XX | XX | | PR | 10-AUG-1999; | 99US-01481711 |
| XX | XX | | PR | 11-AUG-1999; | 99US-01483191 |
| XX | XX | | PR | 11-AUG-1999; | 99US-01483411 |
| XX | XX | | PR | 13-AUG-1999; | 99US-01485651 |
| XX | XX | | PR | 13-AUG-1999; | 99US-01486841 |
| XX | XX | | PR | 16-AUG-1999; | 99US-01493681 |
| XX | XX | | PR | 17-AUG-1999; | 99US-01491751 |
| XX | XX | | PR | 18-AUG-1999; | 99US-01494261 |
| XX | XX | | PR | 20-AUG-1999; | 99US-01497221 |
| XX | XX | | PR | 20-AUG-1999; | 99US-0 |

| | | |
|----|-------------|----------------|
| PR | 16-JUN-1999 | 99US-013946621 |
| PR | 18-JUN-1999 | 99US-013946631 |
| PR | 18-JUN-1999 | 99US-013975050 |
| PR | 18-JUN-1999 | 99US-013976763 |
| PR | 21-JUN-1999 | 99US-013981719 |
| PR | 22-JUN-1999 | 99US-013986999 |
| PR | 23-JUN-1999 | 99US-014005539 |
| PR | 23-JUN-1999 | 99US-014035549 |
| PR | 24-JUN-1999 | 99US-014069599 |
| PR | 28-JUN-1999 | 99US-014082839 |
| PR | 29-JUN-1999 | 99US-014099919 |
| PR | 30-JUN-1999 | 99US-014128799 |
| PR | 01-JUL-1999 | 99US-014184629 |
| PR | 02-JUL-1999 | 99US-014225149 |
| PR | 02-JUL-1999 | 99US-014205559 |
| PR | 06-JUL-1999 | 99US-014232909 |
| PR | 08-JUL-1999 | 99US-014280839 |
| PR | 09-JUL-1999 | 99US-014292909 |
| PR | 12-JUL-1999 | 99US-014297799 |
| PR | 13-JUL-1999 | 99US-014354529 |
| PR | 14-JUL-1999 | 99US-014362649 |
| PR | 15-JUL-1999 | 99US-014400059 |
| PR | 16-JUL-1999 | 99US-014408059 |
| PR | 16-JUL-1999 | 99US-014408069 |
| PR | 19-JUL-1999 | 99US-014432559 |
| PR | 19-JUL-1999 | 99US-014433119 |
| PR | 19-JUL-1999 | 99US-014433329 |
| PR | 19-JUL-1999 | 99US-014434339 |
| PR | 19-JUL-1999 | 99US-014434349 |
| PR | 19-JUL-1999 | 99US-014433359 |
| PR | 20-JUL-1999 | 99US-014435259 |
| PR | 20-JUL-1999 | 99US-014436329 |
| PR | 22-JUL-1999 | 99US-014468849 |
| PR | 22-JUL-1999 | 99US-014484849 |
| PR | 22-JUL-1999 | 99US-014508069 |
| PR | 22-JUL-1999 | 99US-014508089 |
| PR | 22-JUL-1999 | 99US-014508079 |
| PR | 22-JUL-1999 | 99US-014508099 |
| PR | 22-JUL-1999 | 99US-014515129 |
| PR | 23-JUL-1999 | 99US-014514559 |
| PR | 23-JUL-1999 | 99US-014521859 |
| PR | 23-JUL-1999 | 99US-014552249 |
| PR | 26-JUL-1999 | 99US-014572679 |
| PR | 27-JUL-1999 | 99US-014551389 |
| PR | 27-JUL-1999 | 99US-014551819 |
| PR | 27-JUL-1999 | 99US-014581919 |
| PR | 28-JUL-1999 | 99US-014595119 |
| PR | 02-AUG-1999 | 99US-014638689 |
| PR | 02-AUG-1999 | 99US-014638889 |
| PR | 02-AUG-1999 | 99US-014638999 |
| PR | 03-AUG-1999 | 99US-014703889 |
| PR | 04-AUG-1999 | 99US-014720249 |
| PR | 04-AUG-1999 | 99US-014730229 |
| PR | 05-AUG-1999 | 99US-014719229 |
| PR | 05-AUG-1999 | 99US-014726029 |
| PR | 06-AUG-1999 | 99US-014730339 |
| PR | 06-AUG-1999 | 99US-014741619 |
| PR | 09-AUG-1999 | 99US-014749939 |
| PR | 09-AUG-1999 | 99US-014817319 |
| PR | 10-AUG-1999 | 99US-014831199 |
| PR | 11-AUG-1999 | 99US-014831199 |
| PR | 12-AUG-1999 | 99US-014834199 |
| PR | 13-AUG-1999 | 99US-014856559 |
| PR | 13-AUG-1999 | 99US-014868649 |
| PR | 15-AUG-1999 | 99US-014936389 |
| PR | 17-AUG-1999 | 99US-014917559 |
| PR | 18-AUG-1999 | 99US-014942669 |
| PR | 20-AUG-1999 | 99US-014972229 |
| PR | 20-AUG-1999 | 99US-014972339 |
| PR | 22-AUG-1999 | 99US-014992299 |
| PR | 23-AUG-1999 | 99US-014990209 |
| PR | 23-AUG-1999 | 99US-014993029 |

```
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-015138P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155559P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160880P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 10.9%; Score 240.5; DB 3; Length 638;
Best Local Similarity 25.7%; Pred. No. 3e-16;
Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;

```
QY 11 EPPRFLAFGRLLMOVOSRLSSGSSASVYRRCGPNPSGALROFLPPTGTAASAAE 70
DB 249 DPLFLFKVKNCKLQRLAKTSSGSSSEVHKY---SSCTTYALKKTLKGRDYATA---- 301
QY 71 YGFRKRAALQLOQHNRITLVGVFTIHPSPNVPSRCLLLELDVSVSE----- 120
DB 302 YGFCQELGVYKKUKGKNITQL-----IDYEVTDKTLLEQVLNGTMSVKDGRVEDGF 354
QY 121 ---LLYS---SHQGSNM-----MIGHCARDVLEALAFHHEGVHADLK 160
DB 355 IYVAVLEGEIDLAMLSQKWRREIEGSDRTIDENWLRERYOOILQAVVTIHEERIVHSDLK 414
QY 161 PRNIIWSAENECRKLIDFGSPFKGNODVYKTIOTD-----GYRPAEHLQNCIAQGLQ 214
DB 415 PANFL--LVNGFLKLDIFGIA-KAINSDTTNIQRDSOVGLTSLMSPEAFKCNESDENG-- 469
```

```
QY 215 SDPECTSAVDLWSLGIILLEMFSGMKLKHTVRSQGEWKANSATID--HIPAKAVVNAAI 272
DB 470 NTKICGRPSDIWSLACILYQWYTG-RTPFADYKTFW-AKFVITDPNHEITYNQLSNP-- 525
QY 273 PAYHLRDLIKSMTHNDPSRRIRAEEMALCSPPFSIFPAPH----- 311
DB 526 ---WIDLMKKCLADNRQWRIRPELLOHPFLAPPI-PHEPOVKTIKLFSLIAESCGSD 581
QY 312 -----IDLVMLPFPVLRLLNVLDLDYLGNEEYEDVEDYKEEC 351
DB 582 DKANSMISQLEQLNSPAPLPR-----NDVDSRDNQNGQLSRVSELC 624

RESULT 10
AAG38501
ID AAG38501 standard; protein; 720 AA.
XX
AC AAG38501;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47507.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131459P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134570P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135333P.
PR 24-MAY-1999; 99US-013629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
```

PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140654P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148311P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149903P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151067P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156589P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 10.9%; Score 240.5; DB 3; length 720;
Best Local Similarity 25.7%; Pred. No. 3.7e-16;

| | | | | | | | | | |
|-----------|--|--|-----|------------|------|--------|------|------|-----|
| Matches | 105; | Conservative | 57; | Mismatches | 146; | Indels | 101; | Gaps | 17; |
| Qy | 11 | EPPRFLEAFGLWQVSRSLGSSASAYRYRCCNPNPSPALROFLRPPTTGAASAAL | 70 | | | | | | |
| Dd | 331 | DDPLFFKVNKGTLYRLGKIGSGSSEVHKVI---SSDCTIVALKIKLGRDYATA---- | 383 | | | | | | |
| Qy | 71 | YGFKERAAALEOLOGHRNVTVLVGVFTTHFEPNPVPSRCLLHLLDVASE----- | 120 | | | | | | |
| Dd | 384 | YGFCOEIIGYLKKLGKNTIIQL-----IDEVTDTKLVLQEVINGTMSNRGKEGDGF | 436 | | | | | | |
| Qy | 121 | ----LLYS-----SHQGSWM-----MICHCARDVELAALAFLHHGGYHADLK | 160 | | | | | | |
| Dd | 437 | IYMWLEGEIDLAHLMLGSKMRREIGSDRTIDENMLRFYWOOLIAQAVTIHERLYVHADLK | 496 | | | | | | |
| Qy | 161 | PRNLWSANECEFKLIDFGLSFKEGNODVKYIQTD-----GYAPAEALONCIAQAQLO | 214 | | | | | | |
| Dd | 497 | PANFL--LVRGFLKLDIPGIA-KAINSDPTNIQDSQGVLTLSYSPFAFMCNESDENG-- | 551 | | | | | | |
| Qy | 215 | SDTECTSAVDLMISGILLILEMPGSKLKHTRYRSQEMKANSSAIID--HIRASKAVVAAI | 272 | | | | | | |
| Dd | 552 | NTICGRPSDITSLGCILOYWVG-RTPFDYKTFW-AKFKYIDPNHEITYNQLSNP-- | 607 | | | | | | |
| Qy | 273 | PAYHLRDILIKSMLDHPERRIPAEMALCSPPEFSIPAPH----- | 311 | | | | | | |
| Dd | 608 | ---WLIDMKCKCLAMDRNQWRIRPELLQHPLAPP-PHEQVYTKILFSLIASCGSD | 663 | | | | | | |
| Qy | 312 | -----IEDLVMLPTPVLRLLNLVDDVLAGNEBEVEDVEDVEKEC | 351 | | | | | | |
| Dd | 664 | DKANSMTSQLEQLLSNPAPLR-----NDVLDRDQNGQLLSRVSEL | 706 | | | | | | |
| RESULT 11 | | | | | | | | | |
| AAG38500 | ID | AAG38500 standard; protein; 777 AA. | | | | | | | |
| XX | AAG38500; | | | | | | | | |
| XX | 18-OCT-2000 | (first entry) | | | | | | | |
| XX | Arabidopsis thaliana | protein fragment SEQ ID NO: 47506. | | | | | | | |
| XX | Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; Genetic mapping; gene expression control; promoter; termination sequence. | | | | | | | | |
| KW | Arabidopsis thaliana. | | | | | | | | |
| OS | Arabidopsis thaliana. | | | | | | | | |
| XX | EPI033405-A2. | | | | | | | | |
| PD | 06-SEP-2000. | | | | | | | | |
| XX | 25-FEB-2000; 2000EP-00301439. | | | | | | | | |
| PF | 25-FEB-1999; | 99US-0121825P. | | | | | | | |
| XX | 05-MAR-1999; | 99US-0123180P. | | | | | | | |
| PR | 09-MAR-1999; | 99US-0123548P. | | | | | | | |
| PR | 23-MAR-1999; | 99US-0125788P. | | | | | | | |
| PR | 29-MAR-1999; | 99US-0126264P. | | | | | | | |
| PR | 01-APR-1999; | 99US-0126785P. | | | | | | | |
| PR | 06-APR-1999; | 99US-0127462P. | | | | | | | |
| PR | 08-APR-1999; | 99US-0128234P. | | | | | | | |
| PR | 16-APR-1999; | 99US-0128714P. | | | | | | | |
| PR | 19-APR-1999; | 99US-0129845P. | | | | | | | |
| PR | 21-APR-1999; | 99US-0130077P. | | | | | | | |
| PR | 21-APR-1999; | 99US-0130449P. | | | | | | | |
| PR | 23-APR-1999; | 99US-0130510P. | | | | | | | |
| PR | 28-APR-1999; | 99US-0130891P. | | | | | | | |
| PR | 30-APR-1999; | 99US-0131449P. | | | | | | | |
| PR | 30-APR-1999; | 99US-0132048P. | | | | | | | |
| PR | 04-MAY-1999; | 99US-0132407P. | | | | | | | |
| PR | 05-MAY-1999; | 99US-0132484P. | | | | | | | |
| PR | 05-MAY-1999; | 99US-0132485P. | | | | | | | |
| PR | 06-MAY-1999; | 99US-0132486P. | | | | | | | |
| PR | 06-MAY-1999; | 99US-0132487P. | | | | | | | |
| PR | 07-MAY-1999; | 99US-0132488P. | | | | | | | |
| PR | 08-MAY-1999; | 99US-0132489P. | | | | | | | |
| PR | 09-MAY-1999; | 99US-0132490P. | | | | | | | |
| PR | 10-MAY-1999; | 99US-0132491P. | | | | | | | |
| PR | 11-MAY-1999; | 99US-0132492P. | | | | | | | |
| PR | 12-MAY-1999; | 99US-0132493P. | | | | | | | |
| PR | 13-MAY-1999; | 99US-0132494P. | | | | | | | |
| PR | 14-MAY-1999; | 99US-0132495P. | | | | | | | |
| PR | 15-MAY-1999; | 99US-0132496P. | | | | | | | |
| PR | 16-MAY-1999; | 99US-0132497P. | | | | | | | |
| PR | 17-MAY-1999; | 99US-0132498P. | | | | | | | |
| PR | 18-MAY-1999; | 99US-0132499P. | | | | | | | |
| PR | 19-MAY-1999; | 99US-0132500P. | | | | | | | |
| PR | 20-MAY-1999; | 99US-0132501P. | | | | | | | |
| PR | 21-MAY-1999; | 99US-0132502P. | | | | | | | |
| PR | 22-MAY-1999; | 99US-0132503P. | | | | | | | |
| PR | 23-MAY-1999; | 99US-0132504P. | | | | | | | |
| PR | 24-MAY-1999; | 99US-0132505P. | | | | | | | |
| PR | 25-MAY-1999; | 99US-0132506P. | | | | | | | |
| PR | 26-MAY-1999; | 99US-0132507P. | | | | | | | |
| PR | 27-MAY-1999; | 99US-0132508P. | | | | | | | |
| PR | 28-MAY-1999; | 99US-0132509P. | | | | | | | |
| PR | 29-MAY-1999; | 99US-0132510P. | | | | | | | |
| PR | 30-MAY-1999; | 99US-0132511P. | | | | | | | |
| PR | 31-MAY-1999; | 99US-0132512P. | | | | | | | |
| PR | 01-JUN-1999; | 99US-0132513P. | | | | | | | |
| PR | 02-JUN-1999; | 99US-0132514P. | | | | | | | |
| PR | 03-JUN-1999; | 99US-0132515P. | | | | | | | |
| PR | 04-JUN-1999; | 99US-0132516P. | | | | | | | |
| PR | 05-JUN-1999; | 99US-0132517P. | | | | | | | |
| PR | 06-JUN-1999; | 99US-0132518P. | | | | | | | |
| PR | 07-JUN-1999; | 99US-0132519P. | | | | | | | |
| PR | 08-JUN-1999; | 99US-0132520P. | | | | | | | |
| PR | 09-JUN-1999; | 99US-0132521P. | | | | | | | |
| PR | 10-JUN-1999; | 99US-0132522P. | | | | | | | |
| PR | 11-JUN-1999; | 99US-0132523P. | | | | | | | |
| PR | 12-JUN-1999; | 99US-0132524P. | | | | | | | |
| PR | 13-JUN-1999; | 99US-0132525P. | | | | | | | |
| PR | 14-JUN-1999; | 99US-0132526P. | | | | | | | |
| PR | 15-JUN-1999; | 99US-0132527P. | | | | | | | |
| PR | 16-JUN-1999; | 99US-0132528P. | | | | | | | |
| PR | 17-JUN-1999; | 99US-0132529P. | | | | | | | |
| PR | 18-JUN-1999; | 99US-0132530P. | | | | | | | |
| PR | 19-JUN-1999; | 99US-0132531P. | | | | | | | |
| PR | 20-JUN-1999; | 99US-0132532P. | | | | | | | |
| PR | 21-JUN-1999; | 99US-0132533P. | | | | | | | |
| PR | 22-JUN-1999; | 99US-0132534P. | | | | | | | |
| PR | 23-JUN-1999; | 99US-0132535P. | | | | | | | |
| PR | 24-JUN-1999; | 99US-0132536P. | | | | | | | |
| PR | 25-JUN-1999; | 99US-0132537P. | | | | | | | |
| PR | 26-JUN-1999; | 99US-0132538P. | | | | | | | |
| PR | 27-JUN-1999; | 99US-0132539P. | | | | | | | |
| PR | 28-JUN-1999; | 99US-0132540P. | | | | | | | |
| PR | 29-JUN-1999; | 99US-0132541P. | | | | | | | |
| PR | 30-JUN-1999; | 99US-0132542P. | | | | | | | |
| PR | 01-JUL-1999; | 99US-0132543P. | | | | | | | |
| PR | 02-JUL-1999; | 99US-0132544P. | | | | | | | |
| PR | 03-JUL-1999; | 99US-0132545P. | | | | | | | |
| PR | 04-JUL-1999; | 99US-0132546P. | | | | | | | |
| PR | 05-JUL-1999; | 99US-0132547P. | | | | | | | |
| PR | 06-JUL-1999; | 99US-0132548P. | | | | | | | |
| PR | 07-JUL-1999; | 99US-0132549P. | | | | | | | |
| PR | 08-JUL-1999; | 99US-0132550P. | | | | | | | |
| PR | 09-JUL-1999; | 99US-0132551P. | | | | | | | |
| PR | 10-JUL-1999; | 99US-0132552P. | | | | | | | |
| PR | 11-JUL-1999; | 99US-0132553P. | | | | | | | |
| PR | 12-JUL-1999; | 99US-0132554P. | | | | | | | |
| PR | 13-JUL-1999; | 99US-0132555P. | | | | | | | |
| PR | 14-JUL-1999; | 99US-0132556P. | | | | | | | |
| PR | 15-JUL-1999; | 99US-0132557P. | | | | | | | |
| PR | 16-JUL-1999; | 99US-0132558P. | | | | | | | |
| PR | 17-JUL-1999; | 99US-0132559P. | | | | | | | |
| PR | 18-JUL-1999; | 99US-0132560P. | | | | | | | |
| PR | 19-JUL-1999; | 99US-0132561P. | | | | | | | |
| PR | 20-JUL-1999; | 99US-0132562P. | | | | | | | |
| PR | 21-JUL-1999; | 99US-0132563P. | | | | | | | |
| PR | 22-JUL-1999; | 99US-0132564P. | | | | | | | |
| PR | 23-JUL-1999; | 99US-0132565P. | | | | | | | |
| PR | 24-JUL-1999; | 99US-0132566P. | | | | | | | |
| PR | 25-JUL-1999; | 99US-0132567P. | | | | | | | |
| PR | 26-JUL-1999; | 99US-0132568P. | | | | | | | |
| PR | 27-JUL-1999; | 99US-0132569P. | | | | | | | |
| PR | 28-JUL-1999; | 99US-0132570P. | | | | | | | |
| PR | 29-JUL-1999; | 99US-0132571P. | | | | | | | |
| PR | 30-JUL-1999; | 99US-0132572P. | | | | | | | |
| PR | 31-JUL-1999; | 99US-0132573P. | | | | | | | |
| PR | 01-AUG-1999; | 99US-0132574P. | | | | | | | |
| PR | 02-AUG-1999; | 99US-0132575P. | | | | | | | |
| PR | 03-AUG-1999; | 99US-0132576P. | | | | | | | |
| PR | 04-AUG-1999; | 99US-0132577P. | | | | | | | |
| PR | 05-AUG-1999; | 99US-0132578P. | | | | | | | |
| PR | 06-AUG-1999; | 99US-0132579P. | | | | | | | |
| PR | 07-AUG-1999; | 99US-0132580P. | | | | | | | |
| PR | 08-AUG-1999; | 99US-0132581P. | | | | | | | |
| PR | 09-AUG-1999; | 99US-0132582P. | | | | | | | |
| PR | 10-AUG-1999; | 99US-0132583P. | | | | | | | |
| PR | 11-AUG-1999; | 99US-0132584P. | | | | | | | |
| PR | 12-AUG-1999; | 99US-0132585P. | | | | | | | |
| PR | 13-AUG-1999; | 99US-0132586P. | | | | | | | |
| PR | 14-AUG-1999; | 99US-0132587P. | | | | | | | |
| PR | 15-AUG-1999; | 99US-0132588P. | | | | | | | |
| PR | 16-AUG-1999; | 99US-0132589P. | | | | | | | |
| PR | 17-AUG-1999; | 99US-0132590P. | | | | | | | |
| PR | 18-AUG-1999; | 99US-0132591P. | | | | | | | |
| PR | 19-AUG-1999; | 99US-0132592P. | | | | | | | |
| PR | 20-AUG-1999; | 99US-0132593P. | | | | | | | |
| PR | 21-AUG-1999; | 99US-0132594P. | | | | | | | |
| PR | 22-AUG-1999; | 99US-0132595P. | | | | | | | |
| PR | 23-AUG-1999; | 99US-0132596P. | | | | | | | |
| PR | 24-AUG-1999; | 99US-0132597P. | | | | | | | |
| PR | 25-AUG-1999; | 99US-0132598P. | | | | | | | |
| PR | 26-AUG-1999; | 99US-0132599P. | | | | | | | |
| PR | 27-AUG-1999; | 99US-0132600P. | | | | | | | |
| PR | 28-AUG-1999; | 99US-0132601P. | | | | | | | |
| PR | 29-AUG-1999; | 99US-0132602P. | | | | | | | |
| PR | 30-AUG-1999; | 99US-0132603P. | | | | | | | |
| PR | 31-AUG-1999; | 99US-0132604P. | | | | | | | |
| PR | 01-SEP-1999; | 99US-0132605P. | | | | | | | |
| PR | 02-SEP-1999; | 99US-0132606P. | | | | | | | |
| PR | 03-SEP-1999; | 99US-0132607P. | | | | | | | |
| PR | 04-SEP-1999; | 99US-0132608P. | | | | | | | |
| PR | 05-SEP-1999; | 99US-0132609P. | | | | | | | |
| PR | 06-SEP-1999; | 99US-0132610P. | | | | | | | |
| PR | 07-SEP-1999; | 99US-0132611P. | | | | | | | |
| PR | 08-SEP-1999; | 99US-0132612P. | | | | | | | |
| PR | 09-SEP-1999; | 99US-0132613P. | | | | | | | |
| PR | 10-SEP-1999; | 99US-0132614P. | | | | | | | |
| PR | 11-SEP-1999; | 99US-0132615P. | | | | | | | |
| PR | 12-SEP-1999; | 99US-0132616P. | | | | | | | |
| PR | 13-SEP-1999; | 99US-0132617P. | | | | | | | |
| PR | 14-SEP-1999; | 99US-0132618P. | | | | | | | |
| PR | 15-SEP-1999; | 99US-01326 | | | | | | | |

PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-014518P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 03-AUG-1999; 99US-0146389P.
 PR 04-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 05-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 06-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149802P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-015138P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 24-SEP-1999; 99US-0155486P.
 PR 28-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.

PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 10.9%; Score 240.5; DB 3; Length 777;
 Best Local Similarity 25.7%; Pred. No. 4, 2e-16;
 Matches 105; Conservative 57; Mismatches 146; Indels 101; Gaps 17;

QY 11 EPPFLFARGLMOYQSRSGSSASVYVRCCGNPGSPGALROFLPPTGAASAAE 70
 DB 388 DPDLFFKVGKLYQKLGKISGSGSEVHKV---SSDCTIYALKKIKLKGADYAT--- 440
 QY 71 YGFRKRRALFEOQGRNIVTLGVFTIHFSNPVSRCLLELDVSYSE----- 120
 DB 441 YGFCQEIYGLKKLKKKNTIQL-----IDYEVTDKTLQEVNLGTMNKGKRVKEDGF 493
 QY 121 ---LLIYS---SHQCSMW-----MIOHCRADYLAALFLHGGYHADLK 160
 DB 494 IYVLEIYGEIDLAHMLQKWRBIEGSDRTIDENWLRFYWOOLQAVNIHEERIYHSDLK 553
 QY 161 PNIILMSANECEFKLIDPGLSPFEGNODVKYIQTD-----GYRPEALQNCLOAQHQ 214
 DB 554 PANFL--LVKGLKLIIDFGIA-KAINSDTNIQRDSQGTLSYMPKFAFMCNIESENG-- 608
 QY 215 SDPECTSAVDLWLSGIIILEMPSGMLKRTVRSQEMKANSAAID--HIFASKAVVNAI 272
 DB 609 NTIKGGRPSDIMSICILYQWYVG-RTPFADYKTFW-AKFYITDPNHETYNQSNP-- 664
 QY 273 PAYHLRDLIKSMHDDPSRRRIPAEMALCSPPFSIFPAH----- 311
 DB 665 ---WLIDLMKKCLAMDNRQWRIPRILQHPFLAPP1-PHEPOVKTIKLFSLIABSCGSD 720
 QY 312 -----IEDLVNLPVPVLRNLAVLDDVYDGNEEYEDVEDYKESC 351
 DB 721 DKANSMSIQLEQLSNPAPLPR-----NDVIDSRDQOQLSRVSELC 763

RESULT 12
 AAW85028 standard; protein; 544 AA.
 ID AAW85028;
 AC AAW85028;
 XX
 DT 17-OCT-2003 (revised)
 DT 08-FEB-1999 (first entry)
 XX
 DE CDK2-green fluorescent protein usion product.
 XX
 XX Human; CDK2 gene; fusion protein; green fluorescent protein; GFP;
 KW intracellular signalling; chimera.
 XX
 OS Aegourea victoria.
 OS Homo sapiens.
 OS Chimeric.
 OS
 PN WO9845704-A2.
 XX
 PD 15-OCT-1998.
 XX
 XX 07-APR-1998; 98WO-DK000145.
 PF 07-APR-1997; 97DK-00000392.
 XX

XX (NOVO) NOVO-NORDISK AS.
 PA Thastrup O, Petersen Bjorn S, Tullin S, Kasper A, Scudder K;
 PI WPI, 1998-594491/50.
 XX N-PSDB; AA071073.
 DR
 XX Determining effect on signalling pathways in live cells from
 PT redistribution of lumophores - specifically fusions of green
 PT fluorescent protein with a signalling component, and new apparatus,
 PT particularly for identifying toxins and potential therapeutic agents.
 PS
 XX Example 12; Page 208-209; 326pp; English.
 CC The present sequence represents a human CDK2-green fluorescent protein
 CC (GFP) fusion protein. The fusion protein is used in an assay to exemplify
 CC the invention. The specification describes how quantitative information
 CC about the influence of a molecule on a cellular response is obtained by
 CC recording the variation, caused by the molecule, on mechanically intact
 CC living cells, in the spatially distributed light emitted from a
 CC lumophore present in the cells. The variation in light emission is
 CC processed to provide information that correlates spatial distribution to
 CC the degree of the molecule. The method is used to identify agents that
 CC (in)directly affect intracellular signalling, especially to screen for
 CC potential therapeutic agents or toxins, and to identify new drug targets.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 CC
 XX
 SQ Sequence 544 AA;
 Query Match 10.7%; Score 236.5; DB 2; Length 544;
 Best Local Similarity 26.7%; Pred. No. 6.5e-16;
 Matches 111; Conservative 53; Mismatches 154; Indels 97; Gaps 18;
 QY 23 MVOVSRLLGSSSSAVYRVRCNPGSPGALRQF-LPFGTGAASAAYGFRKERALE 81
 DB 4 FQKVEKIGEGTGVVYKAR--NKLGEVVALKKIRLDTEBVPSTAL-----RRTSLK 56
 QY 82 QLOGHNIVTVLVGFTTHFSPVPSRCLLELDVSVSELLYSSHOGCSMMMIQHCARD 141
 DB 57 EL--NHNNIVKLDV--IHTENKL---YLVEFLHODLKKFMDSALVTGIPLEIKSYLFQ 110
 QY 142 VLEALAFHHEGVYADLKPRNIIWSAENECFKLIDFGLSFKGNDVKY---IQTDGVR 198
 DB 111 LLOGLAFCFCHSHRVLRDLKPNLLINAEQS-AIKLADFGIARAFGVPRVRYTHEVVTLMWR 169
 QY 199 APEAEIQLNCLAQAGLSPTCTSAVDLWSLGIILLEMFGSMKLTHTVRSQEWKANSATI 258
 DB 170 APEI-----LGSKYSTAVDWSLGCIFAEW-----VTRALFPGDSE--I 209
 QY 259 DHIFASKAVVN-----AAIPAYHL-----RDLIKSMIHD 287
 DB 210 DQLFRIFRTLTGTPDEVVWPGVTSMDDYKPSFKMARQDFSKVVPPLDEBGRSLLSQMLHY 269
 QY 288 DSRRIIPAEMLCSPPFSIPFAHIEDVLMPTPVRLNLVDDDYLGNEEVEDVVEBV 347
 DB 270 DPKRISAKAALAHPPF-----QP-VTKVPVPHRLMD-----PPVATWV 307
 QY 348 KERCCKYGPVSVLSLVKENGPGQVPEVYANAGDSKAQKLLTGRMFDEKFFVAT 402
 DB 308 SKGEELFTGVPIVLBELGDVNGHKF-SVSGEGSGDATTYCKLT-----LKFICTT 356
 RESULT 13
 ADD21398
 ID ADD21398 standard; protein; 298 AA.
 XX
 AC ADD21398;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Golden hamster cyclin dependent kinase (cdk) 2 protein.
 XX

KW continual growth; cultured cell; cyclin dependent kinase; cdk4; cdk2;
 KW cdk6; activating mutation; cell growth; cell division; cell cycle;
 KW cancer-causing agent; continual growth-induced cell; enzyme;
 KW golden hamster.
 XX
 OS Mesocricetus auratus.
 XX
 PN WO2003044169-A2.
 XX
 PD 30-MAY-2003.
 XX
 PF 15-NOV-2002; 2002WO-US036729.
 XX
 PR 15-NOV-2001; 2001US-0334760P.
 XX
 PA (UTEM) UNIV TEMPLE.
 XX
 PI Reddy PE, Rane SG, Mettus RV;
 XX
 DR WPI; 2003-449813/42.
 XX
 DR N-PSDB; ADD21399.
 XX
 PT A composition for reversibly inducing continual growth in normal cells
 PT comprises a cyclin dependent kinase protein (e.g. cdk4, cdk2 or cdk6) or
 PT its active fragment, derivative, homolog or analog, having an activating
 PT mutation.
 PS
 XX Disclosure; Page 104; 77pp; English.
 CC This invention relates to a novel composition for inducing a reversible
 CC state of a continual growth in cultured cells and comprises at least one
 CC compound comprising a cyclin dependent kinase (cdk)4, cdk2 or cdk6
 CC protein having an activating mutation. Growth and division of living
 CC cells involve a regular series of events and processes that comprise the
 CC cell cycle. Cyclin dependent kinases cdk2, cdk4 and cdk6 are involved in
 CC the control of G1, the point at which cells irreversibly commit to DNA
 CC synthesis and thus enter the cell cycle. The invention is useful in
 CC reversibly inducing continual growth in normal cells and may allow the
 CC screening of cancer-causing agents with the continual growth-induced
 CC cells. The present sequence is that of the golden hamster cdk2 protein
 CC related to the invention. Note: Due to an error in the specification or
 CC sequence listing, the Seq ID numbers given in the disclosure do not
 CC correspond to those given in the sequence listing. It is therefore
 CC unclear which Seq ID number corresponds to which sequence and exactly
 CC which sequence is being claimed.
 CC
 XX
 SQ Sequence 298 AA;
 Query Match 10.7%; Score 235; DB 7; Length 298;
 Best Local Similarity 29.0%; Pred. No. 3.7e-16;
 Matches 98; Conservative 45; Mismatches 117; Indels 78; Gaps 16;
 QY 23 MVOVSRLLGSSSSAVYRVRCNPGSPGALRQF-LPFGTGAASAAYGFRKERALE 81
 DB 4 FQKVEKIGEGTGVVYKAR--NKLGEVVALKKIRLDTEBVPSTAL-----RRTSLK 56
 QY 82 QLOGHNIVTVLVGFTTHFSPVPSRCLLELDVSVSELLYSSHOGCSMMMIQHCARD 141
 DB 57 EL--NHNNIVKLDV--IHTENKL---YLVEFLHODLKKFMDSALVTGIPLEIKSYLFQ 110
 QY 142 VLEALAFHHEGVYADLKPRNIIWSAENECFKLIDFGLSFKGNDVKY---IQTDGVR 198
 DB 111 LLOGLAFCFCHSHRVLRDLKPNLLINAEQS-IKLADFGIARAFGVPRVRYTHEVVTLMWR 169
 QY 199 APEAEIQLNCLAQAGLSPTCTSAVDLWSLGIILLEMFGSMKLTHTVRSQEWKANSATI 258
 DB 170 APEI-----LGSKYSTAVDWSLGCIFAEW-----VTRALFPGDSE--I 209
 QY 259 DHIFASKAVVN-----AAIPAYHL-----RDLIKSMIHD 287
 DB 210 DQLFRIFRTLTGTPDEVVWPGVTSMDDYKPSFKMARQDFSKVVPPLDEBGRSLLSQMLHY 269
 QY 288 DSRRIIPAEMLCSPPFSIPFAHIEDVLMPTPVRL 325

Db 270 DPWKISAKALAHPPF-----QD-VTKPVPHRL 298

RESULT 14
AAB19094
ID AAB19094 standard; protein; 433 AA.

XX AAB19094;

DT 19-FEB-2001 (first entry)

DE Protein encoded by reporter gene YKL161C from protein kinase C pathway.

KM Reporter gene; ergosterol pathway; protein kinase C pathway;

KM invasive growth pathway; biological pathway.

OS Saccharomyces cerevisiae.

PN WO20058520-A1.

XX 05-OCT-2000.

XX 29-MAR-2000; 2000WO-US008555.

XX 31-MAR-1999; 99US-00282243.

XX (ROSE-) ROSETTA INPHARMATICS INC.

XX Roberts CJ;

XX WPI; 2000-664929/64.

DR N-PSDB; AAA86704.

PT Identifying reporter and target genes for particular biological pathway

PT of interest, useful for drug designing, involves identifying a gene which

PT clusters to geneset associated with biological pathway.

XX Disclosure; Fig 20; 239pp; English.

XX The present sequence is encoded by a reporter gene from the protein

CC kinase C pathway. The specification also describes reporter genes from

CC the ergosterol pathway, and the invasive growth pathway. The genes are

CC identified using methods of the invention. The specification describes

CC methods for identifying a reporter gene or a target gene for a particular

CC biological pathway in a cell. The methods comprise identifying a gene

CC which clusters to a geneset associated with the pathway. The gene

CC identified is necessary for normal function of the pathway. The methods

CC are useful for analysing the activity of biological pathways and in drug

CC designing, drug therapies, or other biological agents e.g. insecticides,

CC herbicides, fungicides, antibodies or antivirals. The reporter gene is

CC useful as a specific target for drugs which are designed to enhance,

CC inhibit or modulate a particular biological pathway

XX Sequence 433 AA:

Query Match 10.6%; Score 234.5; DB 3; Length 433;

Best Local Similarity 29.5%; Pred. No. 7,4e-16;

Matches 89; Conservative 45; Mismatches 97; Indels 71; Gaps 16;

Db 226 KDYVDHLNIIQLIGTPEETLQE--IASQKYNYIFQGNIPGRSFESILPGANP--EA 291

QY 278 RDLIKSMHDDPSRRIRPAMALCSFPFSI-----PFA-----PHIDLVWLPTPV 322

Db 292 IELLKRMLEFPDKRITVEDALEHPYLSMMHDIDEEFSCQKTFRPFEPHISMALGNEV 351

QY 323 LR 324

Db 352 IK 353

RESULT 15

ADP87472

ID ADP87472 standard; protein; 433 AA.

XX ADP87472;

DT 23-SEP-2004 (first entry)

DE S cerevisiae glucan synthase pathway protein YKL161C SegID14.

XX glucan synthase pathway; RNA expression; protein expression; YOL1 13W;

KM SKM1; YNR066C; YLR121C; YPS3; YHR209W; YKL161C; YFR030W; FLO1; YFR030W;

XX MET10; antifungal agent.

OS Saccharomyces cerevisiae.

PN WO2004057033-A1.

XX 08-JUL-2004.

XX 17-DEC-2003; 2003WO-US040532.

XX 19-DEC-2002; 2002US-00324035.

XX (ROSE-) ROSETTA INPHARMATICS LLC.

XX Phillips JW;

XX WPI; 2004-500308/47.

DR N-PSDB; ADP87471.

PT Determining whether a molecule affects the function or activity of a

PT glucan synthase pathway in a S. cerevisiae cell by determining a change

PT in the RNA expression or protein expression in the cell of at least one

PT target polynucleotide.

XX Disclosure; SEQ ID NO 14; 132pp; English.

XX This invention relates to a novel method of determining if a molecule

CC affects the function or activity of a glucan synthase pathway in a

CC Saccharomyces cerevisiae cell which comprises determining if the RNA

CC expression or protein expression in the cell of at least one target

CC polynucleotide sequence is changed relative to the expression of said

CC target polynucleotide sequence in the absence of the molecule. The target

CC polynucleotide sequence is selected from the group consisting of YOL1 13W

CC (SKM1), YNR066C, YLR121C (YPS3), YHR209W, YKL161C, YFR030W (FLO1) and

CC YFR030W (MET10). The method is useful for determining whether a molecule

CC affects the function or activity of a glucan synthase pathway in an S

CC cerevisiae cell, possibly allowing development of antifungal agents for

CC use against a variety of pathogens. The present sequence is that of the

CC protein encoded by an S cerevisiae gene which may be used in the method

CC of the invention.

XX Sequence 433 AA:

Query Match 10.6%; Score 234.5; DB 8; Length 433;

Best Local Similarity 29.5%; Pred. No. 7,4e-16;

Matches 89; Conservative 45; Mismatches 97; Indels 71; Gaps 16;

QY 75 KERAALEOQGRNIVTVYGVFTIHFSNVPSRCLL--ELLDVSVSELLYSSHQGCSM 132

```
Db      71 REIKLRLRHGHEINVLFDOT-DIVFYPNGALNGVYLVEELMECDLSQII--RSEQRLD 127
QY      133 WMIOHCARDVLEALAFIHHGYYVHADLKPRNIILMSAENEC-FKLIDFGLS-----FKEG 185
Db      128 AHFQSFITYQLCALKYIHSANVHCDLKPKNLL--VNSDQOLKICNFGLSCSYSENHXYN 185
QY      186 NODVK-YIQTDGYRAPEAEIQNCLAQAGLQSDTECTSAVDLMSGLIILE-----MFSG 238
Db      186 DGFIKGYITSIMYKAPDI-----LITNYQECTKAVDIWSTGCI LAELGRRKPMFEG 235
QY      239 M-----KLKHTVR-----SQEWKANSALIDHTF-----ASKAVVNNAIPAYHL 277
Db      236 KDYVDHLNHLQILGTPEETIQE--IASQKVYNYLFGGNIPGRSFESTILPGANP--EA 291
QY      278 RDLIKSMLHDDPSRRIPAEWALCSPFESI-----PFA-----PHIEDLVMLPTPV 322
Db      292 LELKKMLERDPKKRITVEDALEHPYLSMWMDIDEEFSCQKTRPFEFHEISMALGNEV 351
QY      323 LR 324
Db      352 IK 353
```

Search completed: November 29, 2004, 11:30:38
Job time : 111.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 29, 2004, 11:25:02 ; Search time 4563 Seconds

(without alignments)
4342.403 Million cell updates/sec

Title: US-10-798-532-4

Perfect score: 2206

Sequence: 1 MAGSCAMGABPPRFLEARG.....VATFVPLSAVKRGLYQLTL 419

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 segs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame_plus_p2n.model -DBV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10798532/runat_29112004_112452_15890/app_query.fasta_1.1166
-DB=GenEmbl -QFMT=fastap -SUFFIX=xge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MIN=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10798532.@CNCN_1_1_6317@runat_29112004_112452_15890 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -KONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.1.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_ov.*
7: gb_ov.*
8: gb_ov.*
9: gb_ov.*
10: gb_ov.*
11: gb_ov.*
12: gb_ov.*
13: gb_ov.*
14: gb_ov.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 2195 | 99.5 | 1260 | 6 | CQ715025 Sequence |
| 2 | 2195 | 99.5 | 1260 | 6 | AX262521 Sequence |
| 3 | 2195 | 99.5 | 2598 | 6 | AX262519 Sequence |
| 4 | 2195 | 99.5 | 2909 | 9 | HSAS36197 Homo sapi |

| | | | | | |
|----|--------|------|--------|----|-------------------|
| 5 | 2195 | 99.5 | 4065 | 6 | CQ412057 Sequence |
| 6 | 2195 | 99.5 | 4065 | 6 | AX188296 Sequence |
| 7 | 2188 | 99.2 | 1260 | 6 | AX166543 Sequence |
| 8 | 2188 | 99.2 | 1824 | 6 | AX680136 Sequence |
| 9 | 2183 | 99.0 | 1633 | 10 | RNKISRNA |
| 10 | 2183 | 99.0 | 3244 | 10 | RNU70372 |
| 11 | 2182 | 98.9 | 2041 | 10 | BC058732 |
| 12 | 2174 | 98.5 | 110000 | 2 | AC129141_1 |
| 13 | 2157 | 97.8 | 1254 | 10 | MMPKIS |
| 14 | 1990 | 90.2 | 1367 | 9 | BC026046 |
| 15 | 1890.5 | 85.7 | 2340 | 9 | HSW805327 |
| 16 | 1761.5 | 79.9 | 1157 | 5 | CR386837 |
| 17 | 1721 | 78.0 | 1135 | 9 | AK058195 |
| 18 | 1634 | 74.1 | 2041 | 9 | BC014917 |
| 19 | 1050 | 47.6 | 259562 | 2 | AC094499 |
| 20 | 1045 | 47.4 | 601 | 10 | MMKISRNA |
| 21 | 905 | 41.0 | 105902 | 9 | ALJ59699 |
| 22 | 905 | 41.0 | 167995 | 2 | AC040901 |
| 23 | 901 | 40.8 | 211812 | 2 | AC027205 |
| 24 | 802.5 | 36.4 | 211635 | 2 | AC112763 |
| 25 | 789 | 35.8 | 529 | 6 | AX186343 |
| 26 | 766.5 | 34.7 | 722 | 6 | BD221413 |
| 27 | 729 | 33.0 | 470 | 6 | AX187546 |
| 28 | 721 | 32.7 | 100207 | 10 | AX180177 |
| 29 | 694.5 | 31.5 | 200095 | 10 | AC121606 |
| 30 | 694.5 | 31.5 | 220806 | 10 | AC123650 |
| 31 | 663 | 30.1 | 462 | 6 | AX184680 |
| 32 | 366 | 16.6 | 60640 | 10 | AX180176 |
| 33 | 366 | 16.6 | 185810 | 2 | AC136052 |
| 34 | 361 | 16.4 | 231 | 6 | BD220137 |
| 35 | 327.5 | 14.8 | 205752 | 2 | AC119431 |
| 36 | 311.5 | 14.1 | 419 | 6 | CQ393016 |
| 37 | 311.5 | 14.1 | 419 | 6 | CQ399411 |
| 38 | 267 | 12.1 | 185810 | 2 | AC136052 |
| 39 | 250.5 | 11.5 | 110000 | 8 | CR38128_04 |
| 40 | 250.5 | 11.4 | 1269 | 10 | AB096078 |
| 41 | 250.5 | 11.4 | 1395 | 10 | AB096079 |
| 42 | 249 | 11.3 | 1065 | 10 | HAMCDK2A |
| 43 | 248.5 | 11.3 | 2809 | 3 | AY064464 |
| 44 | 246.5 | 11.2 | 1400 | 10 | RATCDK2A |
| 45 | 246.5 | 11.2 | 2250 | 10 | BC061832 |

ALIGNMENTS

RESULT 1
LOCUS CQ715025 1260 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 959 from Patent WO02068579.
ACCESSION CQ715025
VERSION CQ715025.1 GI:42275682

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

ORIGIN

Alignment Scores: 2.07e-188 Length: 1260
Pred. No.: 2195.00 Matches: 417
Score:

source
1.1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 959 06-SEP-2002;
PB Corporation (NY) (US)
Location/Qualifiers
1.1260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Percent Similarity: 99.76%
 Best Local Similarity: 99.52%
 Query Match: 99.50%

Conservative: 1
 Mismatches: 1
 Indels: 0
 Gaps: 0

US-10-798-532-4 (1-419) x CQ15025 (1-1260)

QY 1 MetAlaGlySerGlyCySaLaATPGLyAlaGluProProARpPheLeuGluAlaPheGly 20
 DB 1 ATGGCGGGATCCGGCTGCGCTGGGGGCGGAGCCGCGCTTTCTGAGGCGCTTCGGG 60
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB 61 CGGCTGTGGAGGTACAGAGCCGCTGGGTAGCGGCGCTCCGCTCGGAGTTCGGGTT 120
 QY 41 ArgCyGlySerGlyAsnProGlySerProProGlyValAlaLeuArgGlnPheLeuProProGly 60
 DB 121 CGCTGCTGGCGCAACCTGCTGCGCCCGCGCCCTCAAGCACTTTGCGCGCAGGA 180
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
 DB 181 ACCACCGGGGCTCGGCTCTGCGCGCGAGTATGTTTCCGCAAGAGAGGGGCGGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB 241 GAACAGTTGCGAGGTACAGAAACATGTCATTTGTATGAGTGTTCATCAATCCACTT 300
 QY 101 SerProAsnValProSerArgCySLeuLeuLeuGluLeuLeuAspValSerValSerGly 120
 DB 301 TCTCCAAATGTGCATCACCGCTGCTGTTCTTGAACCTCGAGATGCACTGTTTGGAA 360
 QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCySerMetTrpMetIleGlnHisCySaLaArg 140
 DB 361 TTGCTCTTATATTCAGTCCACAGGGTGTTCATGTGATGATGATACAGATTCGCCGA 420
 QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuVal 160
 DB 421 GATGTTTGGAGGCCCTTGTCTTTCATCAAGAGGCTATGTCCATGCGGACCTTCAA 480
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGlyCyAspPheLeuIleAspPheGlyLeu 180
 DB 481 CCACGTACATATTTGGAGTGCAGAGATGAAATGTTTAACTCATGTGACTTTGACTT 540
 QY 181 SerPheLeuGlnGlyAsnGlnAspValTyrTyrIleGlnThrAspGlyTyrArgAlaPro 200
 DB 541 AGCTTCAAAAGAGCAATCAGATGTAAATATATTCAGACAGCGGATTCGGGCTCCA 600
 QY 201 GluAlaGlnLeuGlnAsnCySLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
 DB 601 GAAGCAGAAATTCGAAATTCCTTGGCCAGGCTGCGCTGCAAGATGATACAGATGACC 660
 QY 221 SerAlaValaAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetIle 240
 DB 661 TCGAGCTTATCTGTGAGACCTTAGAATCATTTTACTGAAATGTTTCTGAGAAATGAAA 720
 QY 241 LeuValHisThrValArgSerGlnGluTyrPlyAlaAsnSerSerAlaIleIleAspHis 260
 DB 721 CTGAACATATCAGATCTCAGAAATGAGAGCAAAAGTTCTGTATATATATGATCAC 780
 QY 261 IlePheAlaSerIleValAlaValaAsnAlaAlaIleProAlaTyrHisIleLeuAspLeu 280
 DB 781 ATATTTCAGCAAGTAAAGAGGTGTAATGCGCAATTCACACTATACCTTAAAGACCTT 840
 QY 281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 841 ATCAAAAGCATGCTCTCATGTATCCAGCAAGAAATCTCTGCTGAATGAGCATTTGTC 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 901 AGCCCATTTCTTATGACATTTCTTTTGGCCCTCATATTTAAGATCTGGTCATGCTCCACT 960
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrLeuGlnAsnGlnGluTyr 340

DB 961 CCAGTGTCAAGACTGCTGATGTCTGATGATATCTTGCAGATGCAAGAGAAATAT 1020
 QY 341 GluAspValaValaAspValaGlyGlnGlyGlnGlyTyrGlyProValaValSerLeu 360
 DB 1021 GAAGATGTTGTAAAGATGTAAAGAGAGGTCTCAAAAATATATGACCAAGTGTATCTTA 1080
 QY 361 LeuValProGlyGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
 DB 1081 CTGTGTCCAAGAAATCTGTGCGAGAGCAAGTCTTGTGTGATGTGCAATGCTGCT 1140
 QY 381 AspSerValaAlaGlnLeuLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 DB 1141 GATTCAAAGCTGCGCAGAAATTACTGACTGAAAGATTTGATGGAGATTTGTGTG 1200
 QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
 DB 1201 GCTACATTTACCGCTGATGCTGCTTACAAAGAGGGATATCTGTATCAAACTTGCTT 1257

RESULT 2

AX262521 1260 bp DNA linear PAT 26-OCT-2001
 LOCUS AX262521 Sequence 12 from Patent WO0173050.
 DEFINITION AX262521
 ACCESSION AX262521 GI:16511408
 VERSION AX262521.1 GI:16511408
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Meyers, R.
 TITLE 3714, 16742, 23546, and 13887 novel protein kinase molecules and
 uses therefor
 JOURNAL Patent: WO 0173050-A 12 04-OCT-2001;
 Millennium Pharmaceuticals, Inc. (US)

FEATURES
 source 1..1260
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 2,076-188 Length: 1260
 Score: 2195.00 Matches: 417
 Percent Similarity: 99.76% Conservative: 1
 Best Local Similarity: 99.52% Mismatches: 0
 Query Match: 99.50% Indels: 0
 DB: 6 Gaps: 0

US-10-798-532-4 (1-419) x AX262521 (1-1260)

QY 1 MetAlaGlySerGlyCySaLaATPGLyAlaGluProProARpPheLeuGluAlaPheGly 20
 DB 1 ATGGCGGGATCCGGCTGCGCTGGGGGCGGAGCCGCGCTTTCTGAGGCGCTTCGGG 60
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
 DB 61 CGGCTGTGGAGGTACAGAGCCGCTGGGTAGCGGCTCTCCGCGCTCGGATATCGGGTT 120
 QY 41 ArgCyGlySerGlyAsnProGlySerProProGlyValAlaLeuArgGlnPheLeuProProGly 60
 DB 121 CGCTGCTGGCGCAACCTGCTGCGCCCGCGCCCTCAAGCACTTTTGGCGCAGGA 180
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
 DB 181 ACCACCGGGGCTGCGCTCTGCGCGCGAGTATGTTTCCGCAAGAGAGGGGCGGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
 DB 241 GAACAGTTGCGAGGTACAGAAACATGTCATTTGTATGAGTGTTCATCAATCCACTT 300
 QY 101 SerProAsnValProSerArgCySLeuLeuLeuGluLeuLeuAspValSerValSerGly 120

```
Db 301 TCTCCAAATGGCCATCAGCTGTCTGTTGCTTGAACCTCCTGATGTCAGTGTTCGAA 360
Qy 121 LeuLeuLeuTySerSerHisGlnGlyCySerMetTrpMetCileGlnHisCysAlaArg 140
Db 361 TTGCTCTTATATCCAGTCCAGCGGGTGTTCATGTGATGATGATGATGATGATGATGATG 420
Qy 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyTrpValHisAlaAspLeuVal 160
Db 421 GATGTTTGGAGGCCCTTGTCTTTTCTTCATCATGAGGGCTATGTCATGCGAGCTCAAA 480
Qy 161 ProArgAsnIleLeuTrpSerAlaGluAsnGlyCyPheLeuLeuIleAspPheGlyLeu 180
Db 481 CCACGTAAACATATGAGAGTGCAGAGAAATGATTTTAACTCATTTGACTTGGACCTT 540
Qy 181 SerPheLeuGlnGlyValangInaPheValTyTrpIleGlnThrAspGlyTyTrpAlaPro 200
Db 541 AGCTTCAAAGAGAGGCATCAGATGTAAGTATATTCAGACAGACGGGTATCCGGGCTCCA 600
Qy 201 GluAlaGluLeuGlnAsnCySerLeuAlaGlnAlaGlyLeuGlnSerAspTrpGlyCysThr 220
Db 601 GAAAGCAAAATGGCAAAATGCTTGCCCGAGCTGCGCTGCAGAGTGAATACAGATGATACC 660
Qy 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetCys 240
Db 661 TCAAGCTGTGATCTGTGAGAGCTTAAAGATCATTTTACTGGAATGTTTCTCAGGAATGAAA 720
Qy 241 LeuLeuHisThrValArgSerGlnGluTrpValAlaAsnSerSerAlaIleIleAspHis 260
Db 721 CTGAAACATATCAGATCAGATCAGAGATGAGAAAGCAAGTCTCTATATATGATGATCAC 780
Qy 261 IlePheAlaSerIleValAlaValAlaAlaIleProAlaTrpHisIleuAspAspLeu 280
Db 781 ATATTGCGCAGTAAGAGAGTGTGAATGCGCAATTCACGCTATACCTTAAGAGACCTT 840
Qy 281 IleLeuSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
Db 841 ATCAAAAGCATGCTTCATGATGATCCAGAGAGAAATCTCTGTAAGGACATTTGTGC 960
Qy 301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db 901 AGCCCATTTCTTATGAGATCTCTTTTGGCCCTCATATGAAATGATCGGTCAATGCTTCCACT 960
Qy 321 ProValLeuArgLeuLeuAsnValIleuAspAspAspTrpLeuGlnGlnGlnGlnGly 340
Db 961 CCAAGTCTAAAGCTGCTGAATGTGCTGATGATGATATCTTGAAGATGAAAGAGATAT 1020
Qy 341 GluAspValAlaGlnAspValIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
Db 1021 GAAAGATGTTGTGAAGATGTAAGAGAGAGTGCAGAAATATGAGACCAAGGTATCTCTA 1080
Qy 361 LeuValProIleGlnAspLeuProGlyArgGlyGlnValPheValGlnTyTrpAlaAsnAlaGly 380
Db 1081 CTGTGTTCCAAAGGAAATCTCTGCGAGAGCAAGTCTTGTGATGATGCAAAATGCTGGGT 1140
Qy 381 AspSerIleAlaIleGlnIleLeuLeuThrGlyArgMetPheAspGlyLysPheValAla 400
Db 1141 GATTCCAAAGCTCGCAAAATTAATGATGTAAGAGATGTAAGAGATGTAAGATGTAAG 1200
Qy 401 AlaThrPheTyTrpProLeuSerAlaTyTrpLysArgGlyTyTrpLeuTyTrpIleLeuLeu 419
Db 1201 GCTACATTTCTACCGCTGAGTGCCTACAAAGGGGATATGATATCAAACTTGGCTT 1257

RESULT 3
AX262519 2598 bp DNA linear PAT 26-OCT-2001
LOCUS Sequence 10 from Patent WO0173050.
DEFINITION AX262519
ACCESSION AX262519 GI:16511406
VERSION AX262519.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
1 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS Meyers, R.
TITLES 3714, 16742, 23546, and 13887 novel protein kinase molecules and
uses therefor
JOURNAL Patent: WO 0173050-A 10 04-OCT-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1..2598
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
268..1527
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD10238.1"
/db_xref="GI:16511407"
/translation="MAGSCAMGAPRFLAERLMQVSRSLGSSASVYRACCG
NPGSPGALKQFLPGTTGAASAASAEYGFRRKRALAEOLQGRNIVTLVYGVTTIFSP
NVPSCRLLBELLDVSVSLLYSSHOGSGMMIIOHCAVDLEALAFLEHGGVHADLK
PNILIMSANENRCKLIDPLSPKGNODVXYIOTGVAPRAEAIQNCIAOGLSDTE
CTSAVDLNSLIGIILEMFSGKMLKRTVRSQERKANSNIIIDHIFASKAVVNAIATYH
LMDLIKSMLEHDPSSRIIPAEMLCSPPFSIPAPHIEDLVMLPTVRLNLVLDYIL
ENBEYEDVVEDVKEEQYGPVSLVLPKENPGRGVFEVYANAGDSKAAQKLTGR
MDGKRVATFYPFLSAVXRGYLYQTLL"
ORIGIN
Alignment Scores:
Pred. No.: 5.7e-188 Length: 2598
Score: 2195.00 Matches: 417
Percent Similarity: 99.76% Conservative: 1
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 99.50% Indels: 0
Gaps: 0
US-10-798-532-4 (1-419) x AX262519 (1-2598)
Qy 1 MetAlaGlnSerGlnCysAlaTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20
Db 268 ATGGCGGAGATCCGGCTGCGCCCTGCGGCGGAGCGCCCGCTTCTTGAGAGCTTCGGG 327
Qy 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyTrpVal 40
Db 328 CGGCTGTGCAAGTACAGAGCGGTCTGGGTACCGGCTCTCGGCTCGGTATCGGGT 387
Qy 41 ArgCysCysGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
Db 388 CCTGCTGCGGCAACCTGCTGCGCCCGCGGCTTCACAGATCTTGTGCGGCAAG 447
Qy 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyTrpPheArgIleGlnArgAlaAlaLeu 80
Db 448 ACCACCGGGGCTGCGGCTCTCTCGCGGAGTATGTTTCGCAAGAGAGGGGCGGCTG 507
Qy 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyTrpValPheThrIleHisPhe 100
Db 508 GAAAGTTCAGAGGTACAGAAACATCGTATTTGTATGAGGTGTTTAAATCAATCACTT 567
Qy 101 SerProAsnValProSerArgCysLeuLeuLeuGlnLeuAspValSerValSerGlu 120
Db 568 TCTCCAAAGTGCATACAGCTGCTGTTGCTTGAACCTCTGATGTCAGTGTTCGGAA 627
Qy 121 LeuLeuLeuTySerSerHisGlnGlyCySerMetTrpMetIleGlnHisCysAlaArg 140
Db 628 TTGCTCTTATATCCAGTACCAAGGGTGTTCATGAGATATACAGATGTTGCCCA 687
Qy 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyTrpValHisAlaAspLeuVal 160
Db 688 GATGTTTGGAGGCCCTTGTCTTATCATATGAGGGTATGTCATGCGGACCTCAAA 747
Qy 161 ProArgAsnIleLeuTrpSerAlaGluLeuGlnGlyCysPheLeuLeuIleAspPheGlyLeu 180
Db 748 CCACGTAAACATATGAGAGTGCAGAGATGATGATTTTAACTCATTTGACTTGGACTT 807
```

```

QY 181 SerPheLeuGluGlyAsnGlnAspValIleTyrIleGlnThrAspGlyTyrArgAlaPro 200
Db 808 AGCTTCAAGAGAGCATCGATGTAAATATTATTCAGACAGCGGATTCGGCTCCA 867
QY 201 GUAAGGUGUUGGlnGlnGlnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
Db 868 GAAGCAGAAATGGAAAATGCTGGCCGAGCTGGAGTGAAGTAAAGAAATGATACC 927
QY 221 SerAlaValAspLeuTyrSerLeuGlyIleIleLeuLeuGlnUmeCpPheSerGlyMetLys 240
Db 928 TCAGCTGTTGATCTGTCGAGCCCTAGGAATCATTTTACTGGAATGTTCTCAGGAATGAAA 987
QY 241 LeuLysHisThrValArgSerGlnGluTyrLysAlaAsnSerAlaIleIleAspHis 260
Db 988 CTGAACATATCAGATCGATCTCAGAAATGGAAGCAACAGTTCTGCTATTATTGATCAC 1047
QY 261 IlePheAlaSerIleValIleValAlaAlaIleProAlaTyrHisIleuAspLeu 280
Db 1048 ATATTTGGCCGTAAGACAGTGTGAATGCCGCAATTCAGCTTATCCCTAAGAGACCTT 1107
QY 281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnUmeAlaLeuCys 300
Db 1108 ATCAAAAGCATGCTTCATGATGATCCAGACAGAAATTCCTCTGAATGGCATTTGTGC 1167
QY 301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db 1168 AGCCCATTTCTTATAGCTTCTTTTGGCCCTCATATTGAAGATCTGGTTCATGCTTCCACAT 1227
QY 321 ProValLeuArgLeuLeuAsnValIleuAspAspAspTyrLeuGlyLeuGlnGlyGlyTyr 340
Db 1228 CCGATCTTAAGATGCTGGAATGTGCTGAGATGATATTATCTTAAGATGAAGAGAAATAT 1287
QY 341 GluAspValValGlnAspValIleLysGlnGluCysGlnIleTyrGlyProValIleSerLeu 360
Db 1288 GAAGATCTTCTAAGATGTAAAGAGAGCTCTCAAAATATGACCAAGTGTATCTCTA 1347
QY 361 LeuValProLysGlnAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
Db 1348 CTGTGTCCAAGGAAATCTGTGCAGAGCAAGTCTTGTGATGATGCAATGCTGTGGT 1407
QY 381 AspSerIleAlaIleGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValIle 400
Db 1408 GATTCCAAAGCTCGCAGAAATTAATGACTGTGAAGATGTTTATGAGGAAATTTGTGTG 1467
QY 401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db 1468 GCTACATTCACCGCTGAGTGCCTACAAAGGGGATATCTGATCAAACTTGTCTT 1524

RESULT 4
HSAS36197 2909 bp mRNA linear PRI 11-JAN-2003
LOCUS HSAS36197 Homo sapiens mRNA for Kis protein.
DEFINITION AJ536197
ACCESSION AJ536197.1 GI:27657360
VERSION 1
KEYWORDS Kis protein. (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1. Bieche, I., Manceau, V., Curmi, P. A., Laurendon, I., Lechkar, S.,
Leroy, K., Vaudaud, D., Sobel, A., and Maucuer, A.
Quantitative RT-PCR reveals a ubiquitous but preferentially neural
expression of the Kis gene in rat and human
unpublished
2. (bases 1 to 2909)
JOURNAL
REFERENCE
AUTHORS Maucuer, A.
TITLE Direct Submision
JOURNAL Submitted (08-JAN-2003) Maucuer A., U440, Inserm, IFM, 17 rue du
fer a moulin, 75005 Paris, FRANCE
FEATURES
source
1. .2909
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="RZPD:IMAGp98M213680"
/db_xref="RZPD:IMAGp98M213680"
/clone="1.M.A.G.B. Consortium cloneid 1455548"
/ligase_type="2 pooled tumors (clear cell type)"
1..2909
/gene="KIS"
/gene="KIS"
146..1405
/gene="KIS"
/codon_start=1
/product="KIS protein"
/protein_id="CAD60192.1"
/db_xref="GI:27657361"
/translation="MAGSGCAMGAPPPFLAFLGLWQVGRLSGSSASVYRRCG
NGSPGALKQFLPPTGTAASAEYGFRRKRALBOLQGRNLTVLTYGFTTHFSF
NPSRCLLELDIVSELLIYSSHQGSMNMIQHCARDVLEALFLHREGVHADLK
PNIIAMSNECFKIDGLSPKEKNQOVKXIOTGYRAPEALONCLAOGLSDTE
CTSAVDLMSLGIILEMSKMLKRTYASQERKANSASIIHIFPSKAVVNAAPYAH
LNDLTKSMIDHPSRRITPEMLKSPFSIPAPHIDLVMLPTVLRLLVNLVDDYL
ENESEYEDVEDVEKCEKQYGPVSLVLPKENPGRGVFEVANAAGDSKAAQKLTGR
MDGRFVATFPYPLSAVRGXYQLTL"

ORIGIN

Alignment Scores:
Pred. No.: 6,68e-188 Length: 2909
Score: 2195.00 Matches: 417
Percent Similarity: 99.76% Conservative: 1
Best Local Similarity: 99.52% Mismatches: 1
Query Match: 99.50% Indels: 0
DB: Gaps: 0

US-10-798-532-4 (1-419) x HSAS36197 (1-2909)
QY 1 MetAlaGlySerGlyCysAlaThrGlyAlaGluProProArgPheLeuGlnAlaPheGly 20
Db 146 ATGGCGGGATCCGGCTGCCCTCGGGCCGGAGCCGCCGCTTCTGAGGCCCTTCGGG 205
QY 21 ArgLeuTyrGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db 206 CCGCTGTGGCAGTACAGAGCCGTCTGGGTACGGGCTCCCGCTCGGTATCGGGTT 265
QY 41 ArgCysArgGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
Db 266 CGCTGTGGGCAACCTGTGCTGCGCCCGCGCTCAACAGATCTTGTGCGCCAGGA 325
QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlyGluArgAlaAlaLeu 80
Db 326 ACCACCGGGGCTGCGGCTCTGCGCCCGAGTATGTTCCGAAAGAGGGCGGCGCTG 385
QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrGlyValPheThrIleHisPhe 100
Db 386 GAACAGTTGACAGGCTCACAGAAATCGTGAATTTGTATGAGTGTTAACAATCCACTTT 445
QY 101 SerProAsnValProSerArgCysLeuLeuGlnGluLeuAspValSerValSerGlu 120
Db 446 TCTCCAAATGTCCATCAGCTGTCTGTGCTTGAACCTCGAGATGACAGTGTTCGGA 505
QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTyrMetIleGlnHisCysAlaArg 140
Db 506 TTGCTCTTATATTCAGTCAACAGAGTGTTCATGATGATGATACACATTTGTCCGA 565
QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys 160
Db 566 GATGTTTGAAGGCCCTTCTTCTTCTCATCATGAGGCTATGTCATGCGGACCTCAA 625
QY 161 ProArgAsnIleLeuTyrSerAlaGluLeuGluCysPheLysLeuIleAspPheGlyLeu 180
Db 626 CCACGTAAATATTTGTGAGTGCAGAGATGATGATTTTAACTATGACCTTGTGACTT 685
QY 181 SerPheLysGluGlyAlaAsnGlnAspValIleTyrIleGlnThrAspGlyTyrArgAlaPro 200
Db 686 AGCTTCAAGAGAGCATCGATGTAAATATTATTCAGACAGACGCGTATCGGCTCCA 745

```

| | | | |
|------------------------|---|---|-----------------|
| Qy | 201 | Glul1aGl1uenuGnanCyVleu1aGl1aGl1aGl1uenuGnsrAspThGluCythr | 220 |
| Db | 746 | GAACACAGAAATGCAAAATGCTTG3GCCAGGCTG3CCGACAGAGTAACAGAAATGTACC | 805 |
| Qy | 221 | Ser1aVal1aApLeuTTPserLeuGly1le1leLeuLeuGluMetPheSerGlyMetCys | 240 |
| Db | 806 | TCACCTGTGTACTGTGAGCCCTAGAGATCATTTTACTCGAAATGTTCTCGAGATGAAA | 865 |
| Qy | 241 | Leu1yshiStr1a1aArgSerGlnGluTPr1ySa1aAsnSerSer1a1a1e1leAspHis | 260 |
| Db | 866 | CTGAACAATACAGTCAGATCTCAGGAATGAAGCAACAGTTCTGCTAATATTGATCAC | 925 |
| Qy | 261 | 1lePhe1aSer1ySa1aVal1aAsn1a1a1e1lePro1a1a1yTh1e1leu1aArgAspLeu | 280 |
| Db | 926 | ATATTTTGCAGATTAAGCAGTG3GTGAATCCGCAATTCAGCTTACACTTAAGACCTT | 985 |
| Qy | 281 | 1le1ySerMetLeuHisAspAspProSer1aArg1lePro1aGluMetAlaLeuCys | 300 |
| Db | 986 | ATCAAAACATAGCTTTCATGATGATCCAAAGCAGAGAAATTCCTGCTGAATGGCATTTGC | 1045 |
| Qy | 301 | SerProPhePheSer1leProPhe1aProHis1leGluAspLeuValMetLeuProThr | 320 |
| Db | 1046 | AGCCCATCTTTAGCATCTCTTTGGCCCTCATATTAAGATCTGTGATCTTCCCACT | 1105 |
| Qy | 321 | ProVal1eua1aGluLeuAsnVal1euaAspPheAsp1yTleuGlyAsnGluGluTyr | 340 |
| Db | 1106 | CCAGTGTCTMAACCTGCTGAATGTGTGATGATGATTTCTTGAGAAATGAAGAGAAAT | 1165 |
| Qy | 341 | GluAspVal1a1aGluAspVal1yLeuGluGluCysGln1yTyrGlyProVal1aSerLeu | 360 |
| Db | 1166 | GAGATGTGTGAGAGATGTAAAGAGAGATGTCAAAATTTGACCCAGTGTATCTTA | 1225 |
| Qy | 361 | LeuVal1Pro1ySgl1uAsnProGly1aArg1yGlnVal1PheVal1GluTyr1aAsn1aGly | 380 |
| Db | 1226 | CTGTGTCCAAAGAAAATCTTGCGCAGAGACAAAGTCTTGTGTAGTATGACAAATGCTGT | 1285 |
| Qy | 381 | AspSer1ySa1a1aGln1yLeuLeuThrg1aYarMetPheAspGlyLyPheVal1a1 | 400 |
| Db | 1286 | GATTCCAAAGCTGCGCAGAAATTAATCTGACTGAAGAGATTTGATGGAGATTTGTGTGTG | 1345 |
| Qy | 401 | AlaThrPheTyrProLeuSerAlaTyr1aYsaArg1yTyrLeuTyrGlnThrLeuLeu | 419 |
| Db | 1346 | GCTACATTTCAACCGCTAGTGCTTACAGAGGGAGATTCGTATCAAACTTGCTT | 1402 |
| RESULT 5 | | | |
| CO412057 | | 4065 bp | DNA |
| LOCUS | CO412057 | | linear |
| DEFINITION | Sequence 19128 from Patent WO0170979. | | PAT 23-JAN-2004 |
| ACCSSION | CO412057 | | |
| VERSION | CO412057.1 | GI:41319838 | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | Homo sapiens (human) | | |
| REFERENCE | Homo sapiens | | |
| AUTHORS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| TITLE | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| JOURNAL | Lee, J. and Lillie, J. | | |
| FEATURES | Genes, compositions, kits, and method for identification, assessment, prevention, and therapy of ovarian cancer | | |
| source | Patent: WO 0170979-A 19128 27-SEP-2001; Millennium Pharmaceuticals, Inc. (US) | | |
| ORIGIN | location/Qualifiers | | |
| | 1..4065 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="unassigned DNA" | | |
| | /db_xref="taxon:9606" | | |
| Alignment Scores: | | | |
| Pred. No.: | 1,07e-187 | Length: | 4065 |
| Score: | 2195.00 | Matches: | 417 |
| Percent Similarity: | 99.76% | Conservative: | 1 |
| Best Local Similarity: | 99.52% | Mismatches: | 1 |

| Query Match: | 99.50% | Indels: | 0 |
|---|--------|--|------|
| DB: | 6 | Gaps: | 0 |
| US-10-798-532-4 (1-419) x CQ412057 (1-4065) | | | |
| QY | 1 | MechAgiYserGlyCybSaIATpGlyAlaGUpRoPbArqPheLeuGluAlaPheGly | 20 |
| Db | 268 | ATGGCGGGATCCGGCTCCGCTCGGGGGCGGAGCGCGGCTTTCTGGAGGCGCTTCGGG | 327 |
| QY | 21 | ArgLeuTpgInuAlGInserArqLeuGlySerGlySerSerAlaSerValTYrArVal | 40 |
| Db | 328 | CGGCTGTGGCAGGATACAGAGCCGCTGTGGATAGCGGCTCTCCGCGCTCGGTGTATCGGAGTT | 387 |
| QY | 41 | ArgCyecYecGlyAsnProGlySerProProGlyValAlaLeuArqGInPheLeuProProGly | 60 |
| Db | 388 | CGCTGTCGGCAGACCCGTGGCTGCCCCCGCGGCCCTCAAGACGTTCTTGCCGCAAGAA | 447 |
| QY | 61 | ThrThrGlyAlaAlaAlaSerAlaAlaGlyTyrGlyPheArqGlySGlyArqAlaAlaLeu | 80 |
| Db | 448 | ACCACCGGGGTGGGGCTCTGGCGCGGAGATATGTTTCCGCAAAAGAGGGCGGCGCTG | 507 |
| QY | 81 | GluGluLeuGlnGlyHisArqAsnIleValThrLeuTyrGlyValPheThrIleHisPhe | 100 |
| Db | 508 | GAACAGTTTGACGGGTCCACAGAAACATGTGACTTTGTATGGAGGTTTTACATCCACTTT | 567 |
| QY | 101 | SerProAsnValProSerArqCybLeuLeuGlnLeuLeuLeuAspValSerValSerGlu | 120 |
| Db | 568 | TCTCCAAATGGCCATACGCTGTCTGTGTGGTGAACCTCTGAGTGCATAGTTTCGGA | 627 |
| QY | 121 | LeuLeuLeuTyrSerSerSerHisGlnGlyCybSerMetTyrMetIleGlnHisCybAlaArq | 140 |
| Db | 628 | TTCGCTCTTATATTCAGATCCACAGGGTGTTCATATGGATGATACAGCATTTGGCCGA | 687 |
| QY | 141 | AspValLeuGluAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuLys | 160 |
| Db | 688 | GATGTTTTGGAGGGCCCTTGTCTTTCTTCATCATCAGGGCATATGTCCATCGCGAACCTCAA | 747 |
| QY | 161 | ProArqAsnIleLeuTyrPsrAlaGluAsnGluCybPheLysLeuIleAspPheGlyLeu | 180 |
| Db | 748 | CCAGGTACATATTTGTGAAGTGCAGAAAGAAATGTTTTAACTCATTTGACTTTGGACTT | 807 |
| QY | 181 | SerPheLysGlnGlyAsnGlnAspValIleTyrIleGlnThrAspGlyTYrArqAlaPro | 200 |
| Db | 808 | AGCTTCAAAGAGGCATACAGAGTAAAGTATTCACAGACAGCGGATTCGGGCTCCA | 867 |
| QY | 201 | GluAlaGluLeuGlnAsnCybLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCybThr | 220 |
| Db | 868 | GAACACAAATTTGCCAAATTTGCTTGGCCCGGCTGGCTGCAGATGTATACGAATGTACC | 927 |
| QY | 221 | SerAlaValAspLeuTyrPsrLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetLys | 240 |
| Db | 928 | TCACTGTGTATCTGTGAAGCTTAGAAATCATTTTACTCGAAATGTTCTCAGAAATGAAA | 987 |
| QY | 241 | LeuLysHisThrValArqSerGlnGluTyrLysAlaAsnSerSerAlaIleIleAspHis | 260 |
| Db | 988 | CTGAACAATACAGTCAAGTCTCAGAGATGGAAGCAAAAGTTCTGCTATATTATGATCAC | 1047 |
| QY | 261 | IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArqAspLeu | 280 |
| Db | 1048 | ATATTTTCCAGTAAAGCAGTGGTGAATGGCGCAATTCAGGCTTATCACCTTAAAGACCTT | 1107 |
| QY | 281 | IleLysSerMetLeuHisAspAspProSerArqArgIleProAlaGluMetAlaLeuCyb | 300 |
| Db | 1108 | ATCAAAAGCATGCTTCATGATGATCCAGACAGAAAGATTCCTGCTGAATAGCATTTGGC | 1167 |
| QY | 301 | SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr | 320 |
| Db | 1168 | AGCCCATTTCTTATCATTTCTTTGGCCCTCAATTAAGAAATCGATCATGCTTCCACT | 1227 |
| QY | 321 | ProValLeuArqLeuLeuAsnValLeuAspAspAspTyrLeuGlyAsnGlnGlnGluTyr | 340 |
| Db | 1228 | CCAATGTCTAAGACATCGCTGAATATGCTGTGATATATATTTCTTGGAATAGAAAGAAATAT | 1287 |

| | | | |
|---|---|---|-----------------|
| QY | 341 | GlunbepvYalValGluAepValLylgsluGluCyaglulYpYrgYlProValValSerleu | 360 |
| Db | 1288 | GAAGATGTTGTGAGAGATGTAAAGAGAGGTCAAATAATGAGCCAGTGGATCTCTTA | 1347 |
| QY | 361 | LeuValProLYsGluAepProGlyYArgslYlglnValPheValGluTyAlAAsnaIagly | 380 |
| Db | 1348 | CTTGTTCCAAAGGAAATCTCTGCGAGAGACAAGCTTGTGTAGTATGCAATGCTGGT | 1407 |
| QY | 381 | AepSerLYsAlaIaIaGlnLYleLeuLeuThrgLYlArgMetPheAepGlyLYsPheValVal | 400 |
| Db | 1408 | GATTCCAAAGCTGCGCAGAGAAATTAATGACTGAGAGAGATTTGATGAGGAAAGTTTGTGTG | 1467 |
| QY | 401 | AlaThrPheTYrProLeuSerAlaTYrLYsArgGlyTYrLeuTYrGlnThrLeuLeu | 419 |
| Db | 1468 | GCTACATTCTACCGCTGAGTGCCTACAGAGGGGATATCTGTATCAAACCTTGCTT | 1524 |
| RESULT 6 | | | |
| LOCUS | AX188296 | 4065 bp | DNA |
| DEFINITION | Sequence 3991 from Patent WO0142467. | linear | PAT 06-AUG-2001 |
| ACCESSION | AX188296 | | |
| VERSION | AX188296.1 | | |
| KEYWORDS | GI:15139769 | | |
| SOURCE | | | |
| ORGANISM | Homo sapiens (human) | | |
| REFERENCE | Homo sapiens | | |
| AUTHORS | EuharaYota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; | | |
| TITLE | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. | | |
| JOURNAL | Schlegel, R., Deeds, J., Berger, A. and Zhao, X. | | |
| FEATURES | Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer | | |
| source | Patent: WO 0142467-A 3991 14-VUN-2001; Millennium Predictive Medicine, Inc. (US) | | |
| ORIGIN | Location/Qualifiers | | |
| | 1..4065 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="unassigned DNA" | | |
| | /db_xref="taxon:9606" | | |
| Alignment Scores: | | | |
| Pred. No.: | 1,07e-187 | Length: | 4065 |
| Score: | 2195.00 | Matches: | 417 |
| Percent Similarity: | 99.76% | Conservative: | 1 |
| Best Local Similarity: | 99.52% | Mismatches: | 1 |
| Query Match: | 99.50% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |
| US-10-798-532-4 (1-419) x AX188296 (1-4065) | | | |
| QY | 1 | MetAlaGlySerGlyCYsAlaATrPglYAlaGluProArgPheLeuGluAlaPheGly | 20 |
| Db | 268 | ATGGCGGGATCCGGCTGGCGCTGGGGGGCGGAGCGCGCGTTTCTGGAAGCCTTGGG | 327 |
| QY | 21 | ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArgVal | 40 |
| Db | 328 | CGGCTGTGGCAGGTACAGAGCGTCTGGGTAGCGGCTCTCCGCTCGGTATGGGGTT | 387 |
| QY | 41 | ArgCYsCYsGlyAAsnProGlySerProProGlyAlaAlaArgGlnPheLeuProProGly | 60 |
| Db | 388 | CGGTGCTCGGCAACCTCTGGCTCGCGCCCGCGCCCTCAAGCAATTTTGGCGCAGAGA | 447 |
| QY | 61 | ThrThrgLYAlaAlaAlaSerAlaIaIaGluTYrGlyPheArgLYsGluArgAlaAlaLeu | 80 |
| Db | 448 | ACCAACGGGGCTCGCGCTCTGCGCCGCAAGTATGTTTCCGAAAGAGAGGGCGGCTG | 507 |
| QY | 81 | GluGlnLeuGlnGlyHisArgAsnLLeValThrLeuTYrGlyValPheThrLLeHisPhe | 100 |
| Db | 508 | GAAAGATTGACAGGACACAGAAACATCTGACTTTGTATGAGAGTATTACAATCACTTT | 567 |
| QY | 101 | SerProAAsnValProSerArgCYsLeuLeuLeuGluLeuLeuAAspValSerValSerGlu | 120 |
| Db | 568 | TCTCCAAATGTCACATCAAGCTGTCTGTGCTTGTAATCTCTGATGTCAAGTTCAGTTCGGA | 627 |

| | | | |
|---|------------|---|------|
| QY | 121 | LeuLeuLeuLysrSerSerHisGlnGlyCyseSerMetTrpMetIleGlnHisCysAlaArg | 140 |
| Db | 628 | TTGGCTCTTAATATTCAGTCAACAGAGGTTGTTCCATGTGAGATGAATACAGCATTTGGCCGA | 687 |
| QY | 141 | AspValIleuGlnAlaLeuAlaPheLeuHis6HisGlnGlyTyrValHisAlaAspLeuLys | 160 |
| Db | 688 | GATGTTTTGGAGGCCCTTGCTCTTTCTTCATCATAGAGGCTATGTCCATGGGCACTCAAA | 747 |
| QY | 161 | ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLysLeuIleAspPheGlyLeu | 180 |
| Db | 748 | CCAGTCAACATATGTGTGAGTGCAGAGATGAAAGTTTAAACTCATGTGACTTTGGACTT | 807 |
| QY | 181 | SerPheLysGlnGlyAsnGlnAspValLysTyrIleGlnTrpArgPglTyrArgAlaPro | 200 |
| Db | 808 | AGCTTCAAGAGAGGCATCAGAGATGTAAAGTATATTCAGACAGACGGGTATACGGGCTTCA | 867 |
| QY | 201 | GluAlaGluLeuGlnAAsnCyseLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr | 220 |
| Db | 868 | GAACCAAGATTGCAGAAATTCCTTGCCCAAGCTGGCTGCAGATGATACAGATGTATCC | 927 |
| QY | 221 | SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetCysSerGlyMetLys | 240 |
| Db | 928 | TCAGCTGTGATCTGTGTGAGGCTTGAGATCATTTTACTGGAAGTCTTCAGGATGAA | 987 |
| QY | 241 | LeuLysHisTrpValArgSerGlnGlnTrpValAlaAsnSerSerAlaIleAlaAspHis | 260 |
| Db | 988 | CTGAACATATACATGCAGATCTCAGAAATGGAAGGCAGAACAGTCTTGCTATATTGATCAC | 1047 |
| QY | 261 | IlePheAlaSerLysAlaValValAsnAlaAlaIleProAlaTyrHisAlaLysArgAspLeu | 280 |
| Db | 1048 | ATATTGGCAGTAAAGCAGTGGTGAATGCCCAATTCACGCTATACCTTAAGAGACCTT | 1107 |
| QY | 281 | IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnMetAlaLeuCys | 300 |
| Db | 1108 | ATCAAAAGCATGCTTCATGATGATTCACCAAGCAAGAAATTCCTGTGAAGTGGCATTTGGC | 1167 |
| QY | 301 | SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr | 320 |
| Db | 1168 | AGCCCATTTCTTACATCTTCCTTTGGCCCCCTCATATTGAAGTCTGTGATCTCCACT | 1227 |
| QY | 321 | ProValIleuArgLeuLeuAsnValIleuAspAspArgTyrLeuGlyAsnGlnGlnLysTyr | 340 |
| Db | 1228 | CCAGTGCTAAAGACTGCTGAATGTCTGGATGATATTATCTTGAAGATGAAGAGAAATAT | 1287 |
| QY | 341 | GluAspValValAlaGluAspValLysGlnGlnCysGlnLysTyrArgIleProValSerLeu | 360 |
| Db | 1288 | GAAGATGTTGTAGAAGATGATAAAGAGAGAGTCAAAAATATGCAACAGTGTATCTCTA | 1347 |
| QY | 361 | LeuValAlaProLysGluAsnProGluArgGlyGlnValPheValGluTyrAlaAsnAlaGly | 380 |
| Db | 1348 | CTTGTTCCAAAGGAAATTCCTGGCAGAGAGCAAACTCTTGTTGGATATGCAAACTTGCT | 1407 |
| QY | 381 | AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal | 400 |
| Db | 1408 | GATTCCAAGCTGGCGCAAGAAATTCTGACTCGAAGAGATGTTTGATGGCAACTTTGTTGTG | 1467 |
| QY | 401 | AlaIlePheTyrProLeuSerSerAlaTyrLysArgGlyTyrLeuTyrArgIleThrLeuLeu | 419 |
| Db | 1468 | GCTACATTCTACCCGCTGAGTGCCTACACAGAGGGGATATCTGTATCAAACTTGCTT | 1524 |
| RESULT 7 | | | |
| AX166543 | LOCUS | AX166543 | |
| AX166543 | DEFINITION | Sequence 34 from Patent WO0138503. | |
| AX166543 | ACCESSION | AX166543 | |
| AX166543.1 | VERSION | AX166543.1 | |
| GI:14546888 | KEYWORDS | GI:14546888 | |
| Homo sapiens (human) | SOURCE | Homo sapiens (human) | |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | ORGANISM | Homo sapiens | |
| Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | |
| REFERENCE | | | |
| 1 | | | |

AUTHORS Plozman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R.,
 Flanagan, P. and Clardy, D.S.
 TITLE Novel human protein kinases and protein kinase-like enzymes
 JOURNAL Patent: WO 0138503-A 34 31-MAY-2001;

FEATURES
 source Location/Qualifiers
 1.1260
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

| Pred. No.: | 8.88e-188 | Length: | 1260 |
|------------------------|-----------|---------------|------|
| Score: | 2188.00 | Matches: | 416 |
| Percent Similarity: | 99.52% | Conservative: | 1 |
| Best Local Similarity: | 99.28% | Mismatches: | 2 |
| Query Match: | 99.18% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-798-532-4 (1-419) x AX166543 (1-1260)

QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
 DB 1 ATGGCGGGAATCCGGCTGGCGCTGGGGGCGGAGCCGGCGTTTCTGGAGGCGCTTCGGG 60
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTYRArgVal 40
 DB 61 CGGCTGTGGAGGTACAGAGCCGCTGGGTAGGCGGCGCTCCGCTCGGGTTCGGGTT 120
 QY 41 ArgCysGlyYasnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
 DB 121 CGGTGCTGGCGCAACCTGCTGCGCCCGCGCGCTCAAGAGATTCTTGGCCGAGGA 180
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTYRArgPheArgGlyArgAlaAlaLeu 80
 DB 181 ACCACCGGCGCTCGGCTCTGCGCGCGAGTATGTTCCGCAAGAGAGCGGCGCTG 240
 QY 81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTYRArgValPheThrIleHisPhe 100
 DB 241 GAACAGTTGGAGGGTCACAGAAACATCTGACTTTGATGAGTGTTCMAATCCACTTT 300
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
 DB 301 TCTCCAAATGTGCATCACCCCTGCTGTTCCTTGAACCTCTGAGATGTCAGTTTCGGA 360
 QY 121 LeuLeuLeuTYRSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
 DB 361 TTGCTCTTATATTCAGGTACCGAGGGTGTTCATGATGATGATACAGCATTTGCCCCA 420
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisGlnGlyTYRValHisAlaAspLeuLeu 160
 DB 421 GATGTTTGGAGGCCCTTGTCTTCTTCATCAGAGGGGCTCATGCTGCGGACCTCAAA 480
 QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGluCysPheLeuLeuLeuAspPheGlyLeu 180
 DB 481 CCACGTAACATATTTGGAGTGCAGAGATGAAATGTTTAACTCATTTGACTTGGACTT 540
 QY 181 SerPheLeuGluGlyAsnGlnAspValTYRValIleGlnThrAspGlyTYRArgAlaPro 200
 DB 541 AGCTTCAAAGAGGCAATCAGATGTAAGTATATTCAGACAGACGGGTATCGGGCTCCA 600
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 DB 601 GAAGCAGAAATGGAAATGCTTGGCCCAAGCTGGCTGCAGATGATACAGATATGACC 660
 QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluMetPheSerGlyMetLeu 240
 DB 661 TCGGCTTTCATCTGGAGCCTAGAGATCATTTTACTGAAATGTCTTCAGAAATGAAA 720
 QY 241 LeuLeuHisThrValArgSerGlnGluTrpLeuAlaAsnSerSerAlaIleIleAspHis 260
 DB 721 CTGAAACATTCAGATCTCAGAAATGGAAGCAAAAGTTCTGCTATATATGATCAC 780

QY 261 IlePheAlaSerGlyAlaValAlaAsnAlaAlaIleProAlaTYRHisLeuArgAspLeu 280
 DB 781 ATATTTCGACGTAAGACATGCTGAATGCCGAAATTCACGCTTACCTAGACATCTT 840
 QY 281 IleYSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCys 300
 DB 841 ATCAAAAGCATCTTCATGATGATCCAGACAGAAATTCCTGCTGAATGCAATGCAATTGCC 900
 QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
 DB 901 AGCCCATCTTTAGATTCCTTTGCCCCCTCATATGGAAGATCTGTCATGCTCCACT 960
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspTYRLeuGlyAsnGluGluTYR 340
 DB 961 CCAAGTCTAAGACTGCTGAATGTGCTGATGATGATCTTGAAGATGAGAGAAATAT 1020
 QY 341 GluAspValAlaGluAspValIleGluGlnCysGlnIleTYRArgIleProValIleSerLeu 360
 DB 1021 GAAGATGTTGTGAAGATGTAAGAGAGTGTCAAAATATGACACAGTGTATCTCTA 1080
 QY 361 LeuValProIleGluAsnProGlyArgGlyGlnValPheValGluTYRAlaAsnAlaGly 380
 DB 1081 CTGTGTTCCAAAGGAATCTGCGAGAGCAAGCTTTGTTGATGATGCAATGCTGGT 1140
 QY 381 AspSerValAlaAlaGlnIleLeuLeuThrArgIleArgMetPheAspGlyIlePheValAla 400
 DB 1141 GATTCCAAGCTGCGCAGAAATTAATCTGAGTGAAGAGATTTGATGGAGATTTGTGG 1200
 QY 401 AlaThrPheTYRProLeuSerAlaTYRArgArgIleTYRLeuTYRThrLeuLeu 419
 DB 1201 GCTACATTTACCCGCTGAGTCTTCAAGAGGAGATATCTGTATCAACCTTGGCTT 1257

RESULT 8
 AX680136 1824 bp DNA linear PAT 29-MAR-2003
 LOCUS Sequence 21 from Patent WO0181555.
 DEFINITION AX680136
 ACCESSION AX680136
 VERSION AX680136.1 GI:29369916
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Gururajan, R., Patterson, C., Recipon, S.A., Tribouley, C.M.,
 Hafalia, A., Khan, F.Y., Yue, H., Au-Young, J., Bandman, O.,
 Baughn, M.R., Borowsky, M.L., Burford, N., Burrill, J.D., Elliot, V.S.,
 Gandhi, A.R., Kearney, L., Lal, P., Lu, D.A., Lu, Y., Tang, Y.T.,
 Zingler, K.A., Griffin, J.A., Hillman, J.J., Marcus, G.A., Nguyen, D.B.,
 Policky, J., Ramkumar, J., Thornton, M., Walla, N.K. and Walsh, R.T.
 Human Kinases

TITLE

JOURNAL Patent: WO 0181555-A 21 01-NOV-2001;
 Incyte Genomics, Inc. (US)

FEATURES

source Location/Qualifiers
 1.1824
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 3107952CB1"

ORIGIN

Alignment Scores:

| Pred. No.: | 1.49e-187 | Length: | 1824 |
|------------------------|-----------|---------------|------|
| Score: | 2188.00 | Matches: | 416 |
| Percent Similarity: | 99.52% | Conservative: | 1 |
| Best Local Similarity: | 99.28% | Mismatches: | 2 |
| Query Match: | 99.18% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-798-532-4 (1-419) x AX680136 (1-1824)

QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20

```

Db      169 ATGGCGGGAATCCGCGCTGCGCGCGGCGGAGCCGCGCTTTCTGGAGGCGCTTCGCG
Qy      21 ArgLeuTProGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      229 CGGCTGTGGAGGATCAGAGCCGCTCTGGAGGCGCTTCCTCCGCTCGGATGATTCGGGTT 288
Qy      41 ArgCySeGlyAnPProGlySerProProGlyValAlaLeuArgLInPheLeuProProGly 60
Db      289 CGCTGCTGGCGCAACCTGGCTCGCCGCCCGCGCCCTCAAGACGATTCCTGGCGCGAGG 348
Qy      61 ThrThrGlyValAlaAlaSerAlaAlaGlyTyrGlyPheArgGlySerGlyValAlaLeu 80
Db      349 ACCACCGGGGCTCGGCTCTGCGCGCGAGTATGTTTCGCAAGAGAGGCGGCGCTG 408
Qy      81 GlnGlnLeuGlnGlyHisArgAniLeValThrLeuTyrGlyValPheThrIleHisPhe 100
Db      409 GAACAGTTGCAGGGTCAAGAAACATGTCATTTGTAAGAGTGTTCACATCCACTTT 468
Qy      101 SerProAnValProSerArgCySeLeuLeuGlnLeuLeuAnPValSerValSerGly 120
Db      469 TCTCCAATGTGCATCAGCTGCTGTGTTGCTTGAACCTCTGGATGTCAGTGTTTCGAA 528
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlyCySeSerMetTrpMetIleGlnHisCySeAlaArg 140
Db      529 TTGCTCTTATATTCAGGTCAAGGGGTGTTCATGTGATGATACAGCATTTGTGCCGA 588
Qy      141 AspValLeuGlnValAlaLeuAlaPheLeuHisHisGlnGlyTyrValHisAlaAspLeuLys 160
Db      589 GATGTTTGGAGGCCCTTCTCTTTCTTTCATCATGAGGCTATGATCCATGGCGGACCTCAA 648
Qy      161 ProArgAniLeuLeuTrpSerAlaGluAnGlyCySePheLysLeuIleAspPheGlyLeu 180
Db      649 CCAAGTAACATATTTGGAGTGCAGAGAAATGATTTTAACTCATATGACTTTGGACTT 708
Qy      181 SerPheLysGlnGlyAnGlnAnPValLysTyrIleGlnHisArgPglYrArgAlaPro 200
Db      709 AGCTTCAAAAGAGCAATCAGATGTAAATATATTCAGCAACAGGGTATCGGGCTCCA 768
Qy      201 GlnAlaGlnLeuGlnAnCySeLysAlaGlnAlaGlyLeuGlnSerAspThrGlyCySeThr 220
Db      769 GAAGCAAAATTCGAAATTCCTGGCCGAGGCTGGCTGCAGATCATCAGATATACC 828
Qy      221 SerAlaValaAspLeuTrpSerLeuGlyIleIleLeuLeuGlnLumPheSerGlyMetLys 240
Db      829 TCAGCTGTGATCTGTGAGACCTCAGGAATCATTTTACTGGAAATGTTCTCAGGAATGAAA 888
Qy      241 LeuLysHisThrValArgSerGlnGlnTrpLysAlaAsnSerSerAlaIleIleAspHis 260
Db      889 CTGAAACATACAGTCAGATCTCAGGAATGGAAGCAACAGTTCGTATATTGATCAC 948
Qy      261 IlePheAlaSerLysAlaValAlaAsnAlaAlaIleProAlaTyrHisLeuArgAspLeu 280
Db      949 ATATTTCGCCGTAAGACAGTGTGATGCCGCAATTCAGCTCATCACTTAAGACACTT 1008
Qy      281 IleLysSerMetLeuHisAspAspProSerArgArgIleProAlaGlnLumetalLeuCyS 300
Db      1009 ATCAAAAGCATGCTCATGATGATCCAAAGCAAGAAATTCCTGCTGAATGGCATTTGTGC 1068
Qy      301 SerProPhePheSerIleProPheAlaProHisIleGlnAspLeuValMetLeuProThr 320
Db      1069 AGCCCATTTCTTATGCAATTCCTTTTGGCCCTCATATGMAATCTGGTCAGATCCCACT 1128
Qy      321 ProValLeuArgLeuLeuAnValLeuAspAspAspTyrLeuGlnGlnGlnGlnGlyTyr 340
Db      1129 CCAGTGTAAAGATCTCGTAATGTGCTGATGATGATTTATTTTGAAGATGAAGAGAAATAT 1188
Qy      341 GlnAspValValGlnAspValLysGlnGlnCySeGlnLysTyrGlyProValValSerLeu 360
Db      1189 GAGGATCTGTAGAAAGATGTAAAGAGGAGTGTCAAAATATATGACACAGTGTATCTCTA 1248
Qy      361 LeuValProLysGlnAnPProGlyArgGlyGlnValPheValGlyTyrAlaAnAlaGly 380

```

```

Db      1249 CTTGTTCCAAAGGAATCTCGACAGAGCAAGTCTTTGTTAGTATGCCAAATCTGCT 1308
Qy      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db      1309 GATTCAAAGCTGCGGAGAAATTAATCACTGAAGAGATGTTGATGGGAAGTTTGTG 1368
Qy      401 AlaThrPheTyrProLeuSerAlaTyrLysValArgGlyTyrLeuTyrGlnThrLeu 419
Db      1369 GCTACATCTTACCGCTGAGTCTCTAAGAGGAGATATCTGTATCAAACTTGTCTT 1425

RESULT 9
RANKISRNA
LOCUS      RANKISRNA      1633 bp      mRNA      linear      ROD 08-JUL-2002
DEFINITION R. norvegicus mRNA for KIS protein.
ACCESSION  X98374
VERSION    X98374.1 GI:1403531
KEYWORDS   KIS gene.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

REFERENCE
1
AUTHORS    Maucuer, A., Ozon, S., Manceau, V., Gavet, O., Lawler, S., Curmi, P. and
            Sobel, A.
TITLE       KIS is a protein kinase with an RNA recognition motif
JOURNAL    J. Biol. Chem. 272 (37), 23151-23156 (1997)
MEDLINE    97435279
PUBMED     9287318
REFERENCE   2 (bases 1 to 1633)
AUTHORS    Maucuer, A.
TITLE       Direct Subdivision
SUBMITTED  (10-JUN-1996) A. Maucuer, INSERM U440, 17 rue du Fer
            Moulin, 75005 Paris, FRANCE
FEATURES
     source          1..1633
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /tissue_type="brain"
                     /clone_lib="7"
                     /dev_stage="neonate"
     gene            1..1633
                     /gene="KIS"
     CDS             39..1298
                     /gene="KIS"
                     /codon_start=1
                     /protein_id="CAA67021.1"
                     /db_xref="GI:1403532"
                     /db_xref="SwIsr-ProT:063285"
                     /translation="MAGSGCAGAPPPFLFENFGRLMQVQSRISGSSASVYRVRCG
                     /TPGSPGALKQFLPGTTGAASAAEYFRFRERALBQLOGRNITVTLYGVFTIHESP
                     /NTPSRCLLEILDVSELYLSSHQGSMNIOHCARDYLAFLHHEGVHADLK
                     /PNNILMSABEFCFLIDFGLSFKEGNOVKTIQTQCYAPAELEONCLAOAGLQSDTE
                     /CTSAVDLWSLIIILEMFGMKLRITVDSQEKMASSAIIDHIFASKAVVNAALPAVH
                     /LRDLTXSMIHDDPSRRIPAEMLCSPPFSIPAPHIEDLVMLPTVVLNLVLDLDYL
                     /ENRDEYEDVEDVEKCEOKYGVVSLVLPKPNRGQYFVEVYANAGDSKAQKILTGR
                     /MDFGKFAVATFPVLSAYKRGVLYQTL"

ORIGIN
Alignment Scores:
Pred. No.:      3,61e-187      Length:      1633
Score:          2183.00      Matches:      414
Percent Similarity: 99.52%      Conservative: 3
Best Local Similarity: 98.81%      Mismatches:  2
Query Match:    98.96%      Indels:      0
DB:             10      Gaps:      0

US-10-798-532-4 (1-419) x RANKISRNA (1-1633)
Qy      1 MetAlaGlySerGlyCySeAlaAlaTrpGlyAlaGluProProArgPheLeuGlnAlaPheGly 20

```

```

Db      39 ATGGCCGGGCTCCGGCTCGCGTGGGGCCGAGCCGCCGCTTCTGAGGCTTCGGG 98
Qy      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db      99 CGGCTGTGGAGGTACAGACCCGCTTAGGCGAGGCTCTCGGCGCTCGGTGATCCGGGTG 158
Qy      41 ArgCysCysGlyLysAnpProGlySerProProGlyValAlaLeuArgLysPheLeuProGly 60
Db      159 CGGTGCTGGCGCACTCCAGGCTGGCCCCCGGCGCCCTCAAGCATTCCTCCGCGGA 218
Qy      61 ThrThrGlyValAlaAlaSerAlaAlaGlyTyrGlyPheArgLysGluArgAlaAlaLeu 80
Db      219 ACCACCGGGGCTCGCGCTCGGCTGCGAGTATGTTTCGAAAGAGAGAGCGAGCGCTG 278
Qy      81 GluGluLeuGlnGlnLysArgAniLevalThrLeuTyrGlyValPheThrLleIlePhe 100
Db      279 GAGCAGTGGAGGGTCAAGAGAACATCGTACCTTTAAGCAAGCTTTTACCATACACTTC 338
Qy      101 SerProAnValProSerArgCysLeuLeuGluLeuLeuAnpValSerValSerGlu 120
Db      339 TCTCCAAATGTGCATCACCTCTCTCTCTTGAACCTCTGATGTCACGCTTTCGGA 398
Qy      121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetLleGlnHisCysAlaArg 140
Db      399 TTGCTGTATATATCCAGTCAACAGGGTGTCTCATGTGATGATCCAGCACTGTGCAGA 458
Qy      141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuLys 160
Db      459 GATGTCTTGAGGCGCTTGTCTTTCTTTCACATGAGGGCTATGTCCATGAGACCTTCAG 518
Qy      161 ProArgAniLeuLeuTrpSerAlaGluAnpGluCysPheLysLeuLleAspPheGlyLeu 180
Db      519 CCAGCAAACTCTCTGTGAGTGGCGAGAACGATGCTTTAAGCTTATGACTTTGAGCTC 578
Qy      181 SerPheLysGlnGlyAnpValLysTyrLleGlnThrAspGlyTyrArgAlaPro 200
Db      579 AGCTTCAAGAGGCAATCAGAGCGTGAAGTATATTTCAGACAGCGGTATCCAGGCTCT 638
Qy      201 GluAlaGluLeuGlnAnpCysLeuAlaGlnAlaGlyLeuGlnLeuAspThrGluCysThr 220
Db      639 GAAGCAAACTGAGATTCTTGTGCGCCAGCGCGCTGAGAGTGAATACAGAGTATACC 698
Qy      221 SerAlaValAspLeuTrpSerLeuGlyLleLleLeuLeuGlnMetPheSerGlyMetLys 240
Db      699 TCGAGCTTGTATCTCTGAGGCTTGGAAATCATTTTACTGAAATTTCTTCAGAAATGAA 758
Qy      241 LeuLysHisThrValArgSerGlnGluTyrLysAlaAnpSerSerAlaLleIleAspHis 260
Db      759 CTGAAACATACAGTCAGATCTCAGAGAGTGAAGCAAGCTTCTGCTATTATTGATCAT 818
Qy      261 IlePheAlaSerLysAlaValAlaAlaAlaLleProAlaTyrHisLeuAlaGlyLeu 280
Db      819 ATATTTTGGCAGTAAAGAGGTGGAAAGCGCAATTCACCTCAATACCTCAAGAGACTT 878
Qy      281 IleLysSerMetLeuHisAspAspProSerArgArgLleProAlaGluMetAlaLeuCys 300
Db      879 ATCAAAAGCATGCTTCAATGACGACCCAGAGAGAGATCCCTGCTGAGATGCGCTTGTGC 938
Qy      301 SerProPhePheSerLleProPheAlaProHisLleGluAspLeuValMetLeuProThr 320
Db      939 AGCCCAATTTCTTAGCATTTCTTTTGGCCCTCATATGAAATCATGATGCTTCCAACT 998
Qy      321 ProValLeuArgLeuLeuAnpValLeuAspAspArgTyrLeuGlnGluGluGluTyr 340
Db      999 CCAAGTCTCAGACTCCCTCAATGTGCGAATGATGATCATATCTTGAATTAAGATATAT 1058
Qy      341 GluAspValValGluAspValLysGluGlnCysGlnLysTyrGlyProValValSerLeu 360
Db      1059 GAAGATTTTGTAAAGATGTAAAGAGAGATGTCAAGAAATATGAGCAAGTGTCTTCTCTG 1118
Qy      361 LeuValProLysGluAnpProGlyArgGlyGlnValPheValGluTyrAlaAnpAlaGly 380
Db      1119 CTGTGTTCAAGGAAATCTCTGGCAGAGCAAGTCTTGTGTGATGATCGCAACGCTGGA 1178

```

```

Qy      381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db      1179 GATTCAAAGCTGCTCAGAAATTCCTGACTCGGAGAGATTTTGCAGGAACTTTGTGTG 1238
Qy      401 AlaThrPheTyrProLeuSerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db      1239 GCTACATTTACCGGTGATGCTTCAAGAGGGATATCTTATCAAACTTGTCTT 1295

RESULT 10
RNU70372      3244 bp      mRNA      linear      ROD 08-SEP-1999
LOCUS      Rattus norvegicus PAM COOH-terminal interactor protein 2 mRNA,
DEFINITION      complete cds.
VERSION      U70372.2
KEYWORDS      GI:5821767
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 3244)
Alam,M.R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
and Bipper,B.A.
The Novel Kinase P-CIP2 Interacts with the Cytosolic Routing
Determinants of the Peptide Processing Enzyme Peptidylglycine
alpha-Amidating Monooxygenase
J. Biol. Chem. 274 (1999) In press
3 (bases 1 to 3244)
Alam,R., Caldwell,B.D., Johnson,R.C., Darlington,D.N., Mains,R.E.
and Bipper,B.A.
Direct Substitution
Submitted (11-SEP-1996) Neuroscience, Johns Hopkins University
School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
21205, USA
4 (bases 1 to 3244)
Caldwell,B.D., Darlington,D.N., Penzes,P., Johnson,R.C.,
Bipper,B.A. and Mains,R.E.
Direct Substitution
Submitted (03-SEP-1999) Neuroscience, Johns Hopkins University
School of Medicine, 725 N. Wolfe Street, WBSB 902A, Baltimore, MD
21205, USA
REMARK      Sequence update by submitter
COMMENT      On Sep 3, 1999 this sequence version replaced gi:1698778.
FEATURES
source
CDS
/db_xref="GI:5821768"
/db_xref="taxon:10116"
/function="serine protein kinase"
/notes="P-CIP2"
/codon_start=1
/product="PAM COOH-terminal interactor protein 2"
/translation="MAGSGCAGAGAPPFLAAGRLAQVQSLGSSASVYRVACCG
TPGSPGALKQFLPGTTGAASAAYGRKRRALBOLQGRNVTLYGVTTIFSP
NYPSCILIELIDVUSSELYVSSHQSGMMWICARDVLEALFLHREGVHADLK
PNHILMSABNECFKLIDFGLSRFEGNDVKYIQTGYRPAPELONCLAAQALGSDTE
CTSAVDLWSLGIILLESFGMKLKHIVSQQEKANSSAIIIDHIFASKAVVNAAIPAYH
LBDLIKSMHDDPSRRIPALMALCSPPFSIIPAPRIEDLVMLPTVLRLLANVLDDYL
ENDEYEDVVEDEKEECQKYGPVSLVLPKPNRGQYFVEYANAGDSKAAQKLTGR

```

ORIGIN MPDGFVATFVPLSAVKKGYLYQTLL"

Alignment Scores:

| Pred. No.: | 9.42e-187 | Length: | 3244 |
|------------------------|-----------|---------------|------|
| Score: | 2183.00 | Matches: | 414 |
| Percent Similarity: | 99.52% | Conservative: | 3 |
| Best Local Similarity: | 98.81% | Mismatches: | 2 |
| Query Match: | 98.96% | Indels: | 0 |
| DB: | 10 | Gaps: | 0 |

US-10-798-532-4 (1-419) x RNUT0372 (1-3244)

QY 1 MetAlaGlySerGlyCysAlaATPGLyAlaGluProProArghPheLeuGluAlaPheGly 20
 DB 162 ATGGCGGGGCTCGGGCTGGGGGGCGGAGCGCGCGCTTCCTCGAGAGGCTTCGGG 221
 QY 21 ArgLeuTPGLyAlaGlnSerArgLeuGlySerGlySerSerAlaSerValTYraGVal 40
 DB 222 CGGCTGGCGAGGTACAGAGCCCGCTTAGCGAGCGGCTTCCTCGGCTCGGTGTCGGGGTG 281
 QY 41 ArgCysCysGlyAsnProGlySerProProGlyValAlaLeuArgLInPheLeuProProGly 60
 DB 282 CGCTGCTGCGGCACTCCAGGCTCGCCCCCGCGCCCTCAAGCACTTCCTGCTCCGGA 341
 QY 61 ThrThyGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaAlaLeu 80
 DB 342 ACCACCGGGGCTCGCGCTCGGGCTGGAGTATGTTCCGAAAGAGAGCGAGCGCTG 401
 QY 81 GluGlnLeuGlnGlyHisArgAsnLLeValThrLeuTYrGlyValPheThrLLeHisPhe 100
 DB 402 GACGAGTTCAGAGGTTCACAGAAATCTGTACTTTATACAGAGTCTTTACCATACACTTC 461
 QY 101 SerProAsnValProSerArgCysLeuLeuLeuGluLeuLeuAspValSerValSerGlu 120
 DB 462 TCTCCAAATGTGCCATCACCTGTCTGTCTTGAACCTCTGAGTGTCAAGCTTTGGAA 521
 QY 121 LeuLeuLeuTYrSerSerHisGlnGlyCysSerMetTrpMetLLeGlnHisCysAlaArg 140
 DB 522 TTGCTGTATATTCAGTCCAGGCTCGAGGTGCTTCATGTGATGATCCAGCACTGCGCAG 581
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisGlnGlyTYrValHisAlaAspLeuVal 160
 DB 582 GAGTCTCTGAGGCGCTTGTCTTCTTCCACATGAGGCGTATGTCCATGACAGCTTCAG 641
 QY 161 ProArgAsnLLeuLeuTPSerAlaGluAsnGlyCysPheLeuLeuLeaPheGlyLeu 180
 DB 642 CCAAGAAATCTCTGTGAGTGGCGAGACGAGTGTCTTAAGCTTATGACTTTGAGACTC 701
 QY 181 SerPheGlyGluGlyAsnGlnAspValTYrLLeGlnThrAspGlyTYrArgAlaPro 200
 DB 702 AGCTTCAAGAGGCAATCAGAGCTGAGATATTTACAGACAGCGGATTCAGACTCCT 761
 QY 201 GluAlaGluLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerPheThrGlyCysThr 220
 DB 762 GAAGCAGAACTGAGAAATTCCTTGTCCAGCGCGCTTGAAGGTATACAGAGTGAACC 821
 QY 221 SerAlaValAspLeuTPSerLeuGlyLLeLLeuLeuGluMetPheSerGlyMetLeu 240
 DB 822 TCAGCTGTATCTCTGAGACCTTGGAACTTTTACTGAAATGTTCTTCAGAAATGAAA 881
 QY 241 LeuLysHisThrValArgSerGlnGlyTrpLysAlaAsnSerSerAlaLLeLLeaPheHis 260
 DB 882 CTGAACATATCAGTCAATCTCAGAGATGGAAGCAACAGTTCTGTATATTATGATCAT 941
 QY 261 LLePheAlaSerLysAlaValAlaSerAlaAlaLLeProAlaTYrHisLeuAlaGAspLeu 280
 DB 942 ATATTGTCAGTAAAGAGGTGGAATGCGCAATTCACCTTATCATCACTCAAGACCTT 1001
 QY 281 LLeLysSerMetLeuHisAspAspProSerArgArgLLeProAlaGluMetAlaLeuCys 300
 DB 1002 ATCAAAAGCATGCTTCTATGACGACCCCAAGCAAGAGATCCCTGCTGAGATGCGCTTGTGC 1061

QY 301 SerProPheSerLLeProPheAlaProHisLLeGluAspLeuValMetLeuProThr 320
 DB 1062 AGCCCATCTTATGATCTCTTGTCCCTCATATTAAGATCTGTGATCTTCAACT 1121
 QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspArgTYrLeuGlyAsnGluGluTYr 340
 DB 1122 CCAGTCTAGACTCTCTCAATGTGTGATGATGATCTTTCGAAAGAGATGAAATAT 1181
 QY 341 GluAspValAlaGluAspValLysGluGluCysGlnLysTYrGlyProValValSerLeu 360
 DB 1182 GAAGATCTGTGAAGATTTAAAGAGAGTCTCAGAAATATGACACAGTGTCTCTCG 1241
 QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTYrAlaAsnAlaGly 380
 DB 1242 CTGTGTCGAAAGAAATCTGTGAGAGGCAAGCTTGTGTGATGACGCAAAACCTGGA 1301
 QY 381 AspSerValAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 DB 1302 GATTCGAAAGCTGCTCAGAAATGCTGACTGGAGAGATGTTGACGGAGAGTTGTGTG 1361
 QY 401 AlaThrPheTYrProLeuSerAlaTYrLysArgGlyTYrLeuTYrGlnThrLeuLeu 419
 DB 1362 GCTACATTTTACCGCTGAGTGCCTTACAAAGAGGATATCTTATCAAACTTGCTT 1418

RESULT 11
 BC058732 2041 bp mRNA linear ROD 30-JUN-2004
 LOCUS BC058732.1 GI:37194892
 DEFINITION Mus musculus kinase interacting with leukemia-associated gene (leukathin), mRNA (cDNA clone MGC:44770 IMAGE:6414877), complete cds.

ACCESSION BC058732
 VERSION BC058732.1
 KEYWORDS Mus musculus (house mouse)
 SOURCE MGC.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2041)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Cavaletto, T.L., Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshitsugu, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.B., Scherch, A., Schein, J.E., Jones, S.J., and Matra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE JOURNAL
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 2041)
 Strausberg, R.
 Direct Submissions
 Submitted (24-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>

Contact: bento-soares@uiowa.edu, tom-casavant@uiowa.edu
 Ronaldo M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A.,
 Fishler, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K.,
 Schetz, T., Smith, C., Snit, E., Tack, D., Trout, K., Walters, J.,
 Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/MLN at: <http://image.llnl.gov>

Series: Plate: Row: Column: 0
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6754447.

FEATURES

source

1. 2041
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="MGC:64770 IMAGE:6414877"
 /tissue_type="Brain, pool of enriched mouse brain tissues
 from 13.5, 14.5, 16.5 and 17.5 dpc"
 /clone_id="N1H BMAP_FRO"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"

gene

1. 2041
 /gene="Kist"
 /note="synonym: KIS"
 /db_xref="LocusID:16589"
 /db_xref="MGI:1341908"
 154..1413
 /gene="Kist"
 /product="kinase interacting with leukemia-associated gene
 (statmin)"
 /protein_id="AAH58732.1"
 /db_xref="GI:37194893"
 /db_xref="LocusID:16589"
 /db_xref="MGI:1341908"

CDS

/codon_start=1
 /product="kinase interacting with leukemia-associated gene
 (statmin)"
 /protein_id="AAH58732.1"
 /db_xref="GI:37194893"
 /db_xref="LocusID:16589"
 /db_xref="MGI:1341908"
 /translation="MAGSGCAGWAEPFRLEAFRLMVOYSLGSSSASVYKRVCCG
 TPESPGALQKFLPFGTGAAGAAEYGFKEKRALQLOQHNIYTLGVFTHPS
 NVSRCLLELDVSVSELLYSHOGSCMMIOHCARDVLEALFLHGGVYADLK
 PNILSAENBECKLIDPLSEKGNODVYIOTDGYRAPEALQCLAOAGIOSPE
 CTSADVLSGLITLLEMFSGMKLKHTRGSEMKANSATIDHFAKAYVNAIPIYH
 LKRLKSMHLDDEPKRIIPALMALCSPPFSIPFAVHLEDVLPVLRLLNVDLYL
 ENEDEYEDVEDVEKCECKRYGVPVSLVPEKNGQVVEYANADSDXAQDLTRG
 MEDGKEFVATFYFLSAVYKRGYLTLL"

ORIGIN

Alignment Scores:

| Pred. No.: | 6.07e-187 | Length: | 2041 |
|------------------------|-----------|---------------|------|
| Score: | 2182.00 | Matches: | 414 |
| Percent Similarity: | 99.28% | Conservative: | 2 |
| Best Local Similarity: | 98.81% | Mismatches: | 3 |
| Query Match: | 98.91% | Indels: | 0 |
| DB: | 10 | Gaps: | 0 |

US-10-798-532-4 (1-419) x BC058732 (1-2041)

QY 1 MetAAG1SerG1yCysa1aTpg1yAlaGluProProArpPheLeuGluAlaPheGly 20
 Db 154 ATGGCGGGGTCGGGCTGCGGCGGGGCGGAGCCGCGGCTTCTGAGGCGCTTCGGG 213
 QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyraVal 40
 Db 214 CGGCTGGGCGAGGTCCAGAGCCGCGTGGGCGAGCGGCTCGTGGCTCGGTACCGGGTG 273
 QY 41 ArgCysGlyGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
 Db 274 CGGTGCTGGCGGTACCGGCGGCTCGGCGGCGCTCAAGAGTTCTCGCTCGGCGGA 333
 QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrglyPheArgGlyGluArgAlaAlaLeu 80

Db 334 ACCACGGGGCTGCGGCTCGGCGGCGAGTATGTTCCGAAAGAGGCGGCGCTG 393
 QY 81 GluGlnLeuGlnGlyHisArgAsn1LeValThrLeuTyrglyValPheThr1LeHisPhe 100
 Db 394 GAGCAGTTTCAGAGGTCCAGAGAACTCGACTTATATAGAGTCTTTACCAATCACTTC 453
 QY 101 SerProAsnValProSerArgCysLeuLeuGluGluLeuAspValSerValSerGlu 120
 Db 454 TCTCCCAATGTCCTCATCGCTGTCTGCTGTTGAATCTTGATGATGATGATGATGAT 513
 QY 121 LeuLeuLeuTySerSerHisGlnGlyCysSerMetTrpMet1LeGlnHisCysAlaArg 140
 Db 514 TTGCTTTATATTCAGTATCAGGAGCTCTCATGTGATGATGATGATGATGATGATGAT 573
 QY 141 AspValLeuGluAlaLeuAlaPheLeuHisHisGluGlyTyraValHisAlaAspLeuHis 160
 Db 574 GATGTTCTGAGGCGCTCTCTCTTTCTTCAACATGAGGCGTCAATGCAAGCTCAAA 633
 QY 161 ProArgAsn1LeuTrpSerAlaGluAsnGluCysPheLeuLeu1LeAspPheGlyLeu 180
 Db 634 CACGAAACATCTGTGAGTCCGAGATGAGTGTTTTAACCTTATGATCTTGACTTC 693
 QY 181 SerPheLeuGluGlyAsnGlnAspValTyraVal1LeGlnThrAspGlyTyraArgAlaPro 200
 Db 694 AGCTTCAAGAAAGGCAATCAGAGACGTAAAGTATATTCAGACAGAGCGGTATGCGCTCT 753
 QY 201 GluAlaGluLeuGlnAsnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGluCysThr 220
 Db 754 GAGCAGAGCTCAGAACTGCTGGCCGAGCGGCTGCGAGAGTATCAAGAAATGTACC 813
 QY 221 SerAlaValAspLeuTrpSerLeuGlyLeu1LeuLeuGluGlnMetPheSerGlyMetCys 240
 Db 814 TCAAGCTGTATGTGAGAGCTCGGAATCATTTATCGAAATGTTCTCAGAAATGAAG 873
 QY 241 LeuLysHisThrValArgSerGlnGluTrpLysAlaAsnSerSerAla1LeuLeuAsnHis 260
 Db 874 CTGAAGAACATCAGTATCAGATCTCAGAGATGAGAGGAAACAGTTCTGCTATTTGATCAT 933
 QY 934 ATATTTGCGCAGTAAAGCAGTGTGATGCGCAATTCAGCTTACCTCAGAGACCTT 993
 QY 281 1LeuLysMetLeuHisAspAspProSerArgGlu1LeuAlaGluMetAlaLeuCys 300
 Db 994 ATCAAAAGCATCTCATGATGATGACCGGCGAGAAAGATCCGTGAGATGCGCATTTGTC 1053
 QY 301 SerProPhePheSer1LeuProPheAlaProHis1LeGluAspLeuValMetLeuProThr 320
 Db 1054 AGCCCATTTCTTATGATCTCTTTGCGCTCATATTTGAAGATCTGATGATGATGATGAT 1113
 QY 321 ProValLeuArgLeuLeuAsnValLeuLeuAspAspAspTyraLeuGlyAsnGluGluTyra 340
 Db 1114 CCAGTGTCTCAGGCTCTCATGATGCTGATGATGATGATGATGATGATGATGATGAT 1173
 QY 341 GluAspValValGluAspValValysGluGluCysGlnLysTyrglyProValValSerLeu 360
 Db 1174 GAAGATGTTGTAGAGATGTGAAGAGAGTGTCAAGAAATATGACACAGTGTCTCTCG 1233
 QY 361 LeuValProLysGluAsnProGlyArgGlyGlnValPheValGluTyraAlaAsnAlaGly 380
 Db 1234 CTTGTTCCAAAGAAATCTTGCGAGAGACAAAGCTTCGTTGATGATGATGATGATGATGAT 1293
 QY 381 AspSerLysAlaAlaGlnLysLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
 Db 1294 GATTCCAAAGCTGTCAAGAGTTGCTGACTGGAGAGAGATGTTTGAACGGAAGTTTGTG 1353
 QY 401 AlaThrPheTyraProLeuSerAlaTyraLysArgGlyTyraLeuTyrglyThrLeuLeu 419
 Db 1354 GCTAACATCTACCGCTGAGTGTCTTCAAGAGGAGATCTTTATCAAACTTGCTT 1410
 RESULT 12
 AC129141_1/c

WPCOMMENT
Sequence split into 4 fragments LOCUS AC129141 Accession AC129141
Fragment Name Begin End
AC129141_0 1 110000
AC129141_1 100001 210000
AC129141_2 200001 310000
AC129141_3 300001 363206
Continuation (2 of 4) of AC129141 from base 100001 (AC129141 Rattus norvegicus clone CH2)

Alignment Scores:
Prod. No.: 8 44e-184 Length: 110000
Score: 2174.00 Matches: 413
Percent Similarity: 99.28% Conservative: 3
Best Local Similarity: 98.57% Mismatches: 3
Query Match: 98.55% Indels: 0
DB: Gaps: 0

US-10-798-532-4 (1-419) x AC129141_1 (1-110000)

QY 1 MetAlaGlySerGlyCysAlaTrpGlyAlaGluProProArgPheLeuGluAlaPheGly 20
Db 70316 ATGGCCGGGCTCGCTGCGCTGGGGGCGCGAGCCGCGCTTCTGGAGGCTTCGGG 70257

QY 21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTyrArgVal 40
Db 70256 CGGCTGGGAGGTACAGAGCCGCTTAGGAGGCGGCTCTCGGCTCGGTACCGGGTG 70197

QY 41 ArgCysGlyValAsnProGlySerProProGlyValAlaLeuArgPheLeuProProGly 60
Db 70196 CGCTGCTGGGCACTCCAGGCTGCGCCCGCGGCGCTTCAAGAGTTCCTCCGCGGA 70137

QY 61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrGlyPheArgGlySerGluArgAlaLeu 80
Db 70136 ACCAGCGGGGCTCGGCTCGGCTGGCGAGTATGTTTCGAAAGAGAGGCGAGCGTG 70077

QY 81 GluGlnLeuGlnGlnHisArgAsnIleValThrLeuTyrGlyValaPheTrpIleHisPhe 100
Db 70076 GAGCAGTTGGAGGATCAGAGAACTGTGACTTTATACGAGCTTTACATACACTTC 70017

QY 101 SerProAsnValProSerArgCysLeuLeuGlnLeuLeuAspValSerValSerGlu 120
Db 70016 TCTCCAAATGTGCATACGCTGTCTCTTCTTGAACCTCGATGTCAGCCTTCGGA 69957

QY 121 LeuLeuLeuTyrSerSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
Db 69956 TTGCTGTATATTCCAGTCAACAGGGTGTCTCATGTGATGATACGACTGTGCAGA 69897

QY 141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTyrValHisAlaAspLeuVal 160
Db 69896 GATGTTCTGGAGGCTTGTCTTTCTTACCATGAGGCTATGTCATGCACTCAAA 69837

QY 161 ProArgAsnIleLeuTrpSerAlaGluAsnGlyCysPheLeuIleAspPheGlyLeu 180
Db 69836 CCACGAAACACTCTGTGGAGTGGCGGAGAACGATGCTTTAAGCTTATGACTTC 69777

QY 181 SerPheLeuGlnGlyAsnGlnAspValTyrTyrIleGlnThrAspGlyTyrArgAlaPro 200
Db 69776 AGCTTCAAAAGAGCAATCAGAGCGTGAATATATTCAGACAGCGGTATTCAGACTC 69717

QY 201 GluAlaGlnLeuGlnAsnGlyLeuAlaGlnAlaGlyLeuGlnSerAspTrpGluCysThr 220
Db 69716 GAAGCAGAACTGCAGAATTCCTTGGCCAGCGGCTCGAGAGTATACAGAGTACC 69657

QY 221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnLeuPheSerGlyMetCys 240
Db 69656 TCAAGCTTATATCTCTGAGACCTTGGAAATCATTTTACTGAAATGTCTTCAGAAATGAAA 69597

QY 241 LeuLeuHisThrValArgSerGlnGluTyrPheAlaAsnSerSerAlaIleIleAspHis 260
Db 69596 CTGAAACATACAGTCAAGATCTCAGAGTGAAGCAAAAGTTCTGCTATATATGATCAT 69537

QY 261 IlePheAlaSerValAlaValAlaAsnAlaAlaIleProAlaTyrHisIleLeuArgAspLeu 280

Db 69536 ATATTGGCAATGAAGCACTGTGAATGCCGAATTCACGCTTATCATCAGACACTT 69477

QY 281 IleYsSerMetLeuHisAspAspProSerArgAlaIleProAlaGluMetAlaLeuCys 300
Db 69476 ATCAAAAGCACTCTCATACAGACCCAGCAAGAGATCCCTGTGAGATGCGCTTGTGC 69417

QY 301 SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
Db 69416 AGCCCAATCTTAAAGATTCCTTTGCCCCCTCATATTAAGATCTGTGATGCTTCAACT 69357

QY 321 ProValLeuArgLeuLeuAsnValLeuAspAspTyrLeuGlyValAsnGluGluTyr 340
Db 69356 CAGAGCTCAGACTCTCATATGTCTGATGATGACTATCTTGAATAATCAATGAATAT 69297

QY 341 GluAspValAlaGluAspValTyrGlnGluCysGlnTyrGlyProValValSerLeu 360
Db 69296 GAAGATGTTGTGAAGATGTAAAGAGAGTGTCAGAATATGACCAAGTGGTTCTCTG 69237

QY 361 LeuValProLeuGluAsnProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGly 380
Db 69236 CTGTTCCTAAAGAAATCTGTGAGAGCAAGCTTTGTGAGTATGCGCAAAACCTGGA 69177

QY 381 AspSerValAlaAlaGlnLeuLeuLeuThrGlyArgMetPheAspGlyLysPheValVal 400
Db 69176 GATTCCTAAAGCTGCTCAGAAATGCTGACTGAGAGATGTTGATGGGAAGTTGTTGG 69117

QY 401 AlaThrPheTyrProLeuSerAlaTyrIysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db 69116 GCTACATTTACCCGTGAGTGTCTTACAAAGAGGAGTATCTTATCAAACTTGGCT 69060

RESULT 13
LOCUS MPMKIS 1254 bp mRNA linear ROD 08-JUL-2002
DEFINITION Mus musculus partial mRNA for KIS protein kinase (Kis gene).
ACCESSION Y10725
VERSION Y10725.2 GI:21726712
KEYWORDS Kis gene; protein kinase.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Maucuer A., Ozon S., Manceau V., Gavet O., Lawler S., Curmi P. and Sobel A.
TITLE KIS is a protein kinase with an RNA recognition motif
JOURNAL J. Biol. Chem. 272 (37), 23151-23156 (1997)
MEDLINE 97435279
PUBMED 9287318
REFERENCE
AUTHORS Maucuer A.
TITLE Direct Submision
JOURNAL Submitted (23-JAN-1997) Maucuer A., INSERM U440, U440, 17 Rue du Fer Moutin, Paris, 75005, FRANCE
REMARK
AUTHORS Maucuer A.
TITLE Direct Submision
JOURNAL Submitted (02-JUL-2002) A. Maucuer, INSERM U440, U440, 17 Rue du Fer Moutin, Paris, 75005, FRANCE
COMMENT
On Jul 10, 2002 this sequence version replaced gi:1806129.
Overlaps with related sequences X82320, X10725.
FEATURES
source location/Qualifiers
1..1254
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="X22"
/tissue_type="brain"
/clone_lib="lambda gti10"
/dev_stage="neonate"
1..1254
/gene="Kis"
11..>1254

```

/gene="Kis"
/codon_start=1
/product="Kis protein kinase"
/protein_id="CAA71714.2"
/db_xref="GI:21726713"
/db_xref="GOA:P97343"
/db_xref="Swiss-Prot:P97343"
/translation="MAGSGCAGMAGPRLFAARGLIMQYRSLGSGSSAYVRRCG
TBSPPGALAKQPLPRTGTAAANAAYPRKEAALFQAGNHNITLYIGVPTIRSP
NPFSRLLELDVSVSELIVSSHQSCSRMMIQCARDVLEALATHHGIVHADLK
PNTIWSAENECFKLIDFGISFREGQDVYIQTIDYRAPEALQNCIAQSGQSTE
CTSAVDLMISGIIILEMFSGMKHTVRSQEMKANSALIDHIFASKAVVNAIIPYH
LRILKSMLEHDDPGRIPALMAKSPFESIPEAPHIEDLVMLPTPLRLINLVDDLYL
ENDEVEDYVEDYKRCOKYGPVSLIVPRENREGQVPEVYANADSKAQLTLGR
MPDKRVTVAATFPLSAIKKGYL"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      5,55e-185      Length:      1254
Score:          2157.00      Matches:      409
Percent Similarity: 99.28%      Conservative: 2
Best Local Similarity: 98.79%      Mismatches: 3
Query Match:     97.78%      Indels:      0
DB:              10      Gaps:      0

```

US-10-798-532-4 (1-419) x MMPKIS (1-1254)

```

QY      1 MetAlGlySerGlyCyValATrGlyAlaGluProProArgPheLeuGluAlaPheGly 20
DB      11 ATGGCGGGGTCGGCGCTGCGCGGGGGCGCGAGCCGCGCTTCTCGAGGCGCTTCGGG 70
QY      21 ArgLeuTrpGlnValGlnSerArgLeuGlySerGlySerSerAlaSerValTYrArgVal 40
DB      71 CGGCTGGGAGGTCAGAGCCGCTGGGAGGCGGCTGTCGGCGCTGCGGTACCGGGTG 130
QY      41 ArgCySerGlyAsnProGlySerProProGlyAlaLeuArgGlnPheLeuProProGly 60
DB      131 CGGTGCTGGCGTACC CGGGCTCGCCCGCGCGCCCTCAAGAGATTCCGCTCGCGGA 190
QY      61 ThrThrGlyAlaAlaAlaSerAlaAlaGluTYrGlyPheArgGlyGluArgAlaLeu 80
DB      191 ACCACCGGGGCTCGCGCTCGCGCGCGAGTATGTTCCGCAAGAGAGCGCGGCTG 250
QY      81 GluGlnLeuGlnGlyHisArgAsnIleValThrLeuTYrGlyValPheThrIleHisPhe 100
DB      251 GAGCAGATGGAGGCTCAAGAGAACATGTGACTTATATGAAGCTTTACCATACACTTC 310
QY      101 SerProAsnValProSerArgCyLeuLeuLeuGluLeuLeuAspValSerValSerGly 120
DB      311 TCTCCCAATGTGCATACACGCTGTCTGCTTGAACCTCTGATGTGCACTGTTCCGAA 370
QY      121 LeuLeuLeuTYrSerSerHisGlnGlyCySerSerMetTrpMetIleGlnHisCyValArg 140
DB      371 TTGCTTTTATATTCAGATCATCAGGCGCTCTCCATGTGATGATACAGCACTGTGCCGA 430
QY      141 AspValLeuGlnAlaLeuAlaPheLeuHisGlnGlyTYrValHisAlaAspLeuVal 160
DB      431 GATGTTCTGAGGCGCTCGCTTTTTCACATGAGGCGCTACGTCATGACAGACTTCAAA 490
QY      161 ProArgAsnIleLeuTrpSerAlaGluAsnGlyCyAspPheLeuIleAspPheGlyLeu 180
DB      491 CCAGCAAACTCTCTGTGAGTGCAGAGAACATGATTTTAAGCTTATGACTTTGACATC 550
QY      181 SerPheLeuGluGlyAsnGlnAspValTYrIleGlnThrAspGlyTYrArgAlaPro 200
DB      551 AGCTTCAAGAGAGCATCAGACGTAAGTATATTCAGACAGACGGGTTTCCGCGCTCT 610
QY      201 GluAlaGluLeuGlnAsnCyLeuAlaGlnAlaGlyLeuLeuIleAspThrGlnTyThr 220
DB      611 GAAGCAGAGCTGAGAACTCTTGGCCAGGCGGCTGAGAGATGATACAGATGATACC 670
QY      221 SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGluLeuPheSerGlyMetTy 240

```

```

DB      671 TCAGCTGTGATCTGTGAGAGCTCCGGAATCATTTTACTGGAATGTCTCAGAAATGAG 730
QY      241 LeuValHisIleThrValArgSerGlnGluTrpValAsnSerSerAlaIleIleAspHis 260
DB      731 CTGAAACATACAGTCAGATCTCAGGAATGAAAGGCAACAGTTCTGCTATTTATGATCAT 790
QY      261 IlePheAlaSerLeuAlaValAlaAsnAlaIleIleProAlaTYrHisIleuArgAspLeu 280
DB      791 ATTTTTCAGATTAAGACATGTGTGAATGCCGCAATTCAGCTTATTCACCTCAGAGACTT 850
QY      281 IleValSerMetLeuHisAspAspProSerArgArgIleProAlaGluMetAlaLeuCy 300
DB      851 ATCAAAAGCATGCTTCATGATGACCCGGCAGAGAAATCCCTGTGATGATGACATGTGCG 910
QY      301 SerProPhePheSerIleIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB      911 AGCCCATTTCTTAAAGCATTTCTTTTGGCCCTCATATGGAAGATCTGTGTATGCTCGACT 970
QY      321 ProValLeuArgLeuLeuAsnValLeuAspAspAspTYrLeuGlyAsnGluGluTYr 340
DB      971 CAGTCTCAGGCTCTCTCATGTGCTGATGATGATGATTAACCTTGAAATGAAATGAAATAT 1030
QY      341 GluAspValValGluAspValIleGluGluGlyCyGlnIleTYrGlyProValValSerLeu 360
DB      1031 GAAGATGTTGTGAAGATGTGAAGAGGAGTGTCAAGAAATATGACCACTGGTTCTCTG 1090
QY      361 LeuValProIleGluAsnProGlyArgGlyValIlePheValGluTrpAlaAsnAlaGly 380
DB      1091 CTGTTCCTCAAGAGAAATCTGTGCAAGAGCAAGATCTTGTTGATGACCGAAGCTGTGT 1150
QY      381 AspSerLeuAlaAlaGlnIleLeuLeuThrGlyArgMetPheAspGlyAspPheValVal 400
DB      1151 GATTCCAAGCTGCTCAGAAAGTTGCTGACTGGAGAGATTTTGACCGGAAGTTGTGTG 1210
QY      401 AlaThrPheTYrProLeuSerAlaTYrIleAspArgIleTYrLeu 414
DB      1211 GCTACATTTACCCGCTGAGTGCCTACAGAGGGGATATCTT 1252

RESULT 14
BC026046      1367 bp      mRNA      linear      PRI 30-JUN-2004
LOCUS      BC026046.1      GI:19684094
KEYWORDS      MGC.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE
AUTHORS      Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Scheffer,C.F.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmitz,J., Myers,R.M.,
Butterfield,Y.S., Kizylinski,M.I., Skalka,U., Smalins,D.B.,
Scherer,A., Schein,J.E., Jones,S.J., and Matra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

```

JOURNAL

PUBMED 12477932
2 (bases 1 to 1367)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
infobgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu, Parvaneh Saeedi, JR Santos, Angelique Scherch, Ursula Skalska, Duane Smalins, Jeff Scott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 42 Row: f Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21389438.
Location/Qualifiers
1. 1367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:39391 IMAGE:4649261"
/issue_type="Eye, retinoblastoma"
/clone_lib="NIH MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOT87"
1. 1367
/gene="KIS"
/db_xref="LocusID:127933"
159. 1193
/gene="KIS"
/codon_start=1
/product="kinase interacting stathmin"
/protein_id="AAH26046.1"
/db_xref="GI:19684095"
/translaton="MAGSGCAGBAPPRFLFAFRMLQVOSRLSGSSASVYVRCCG
NPGSPGALKOPLPGCTGGAASAAYGPFKEKRALEOLQGHNTITLVGPTIHSP
NYSRCLLELDIVSYSELLYSHGCSKMTQHCARDYLEALILHGGYHDLK
PRNLTWSAENECFKLIDFGLSFEKNQDVYTIQTDYRAVEALQCLQAGLSQSTE
CTSAVDLWSIGLILMFSGMKLHTVRSGEMNSATIDHIFAKAVVNAIPYVH
LRLLKSMHLDPSRRIPAMALCSPFSPFAPHIEDIVMLPTPLRLINLVDDVYL
ENEEVEGIC"
ORIGIN
Alignment Scores:
Pred. No.: 7.45e-170 Length: 1367
Score: 1990.00 Matches: 385
Percent Similarity: 92.14% Conservative: 2
Best Local Similarity: 91.67% Mismatches: 2
Query Match: 90.21% Indels: 31
DB: 9 Gaps: 1
US-10-798-532-4 (1-419) x BC026046 (1-1367)
QY 1 MetAAGTGCserGlyCysAlaATTPGlyAlaGluProProArgRphelengluAlaIaphegLy 20
DB 159 ATGGCGGAGATCCGAGCTGGCGCGCGAGCCGCCGCTTTCTGGAGGCGCTTCGGG 218

QY ArgLeuTrpGluValGlnSerArgLeuGlySerSerAlaSerValTyraArgVal 40
DB CGGCTGTGGCAGATACAGAGCCGTCGGGTACGGGCTCTCCGCTCGGTATCGGGTT 278
QY ArgCysCysGlyAsnProGlySerProProGlyAlaAlaArgGlnPheLeuProProGly 60
DB CGCTGCTGGGGAACCTGTGCTGCCCCCGCGCCCTTCAACAGTTCTTGCCGCGAGA 338
QY ThrThrGlyAlaAlaAlaSerAlaAlaGluTyrglyPheArgLyGluArgAlaAlaLeu 80
DB ACCACCGGGGCTGCGGCTCTGCGCCGAGTATGTTCCGAAAGAGGGCGCGCTG 398
QY GlnGlnLeuGlnGlyHisArgAsnIleValThrLeuTyrglyValPheThrIleHisPhe 100
DB GAACAGTTCAGAGGTACAGAAACATCGACCTTGATGAGGTGTTCAATCCACTTT 458
QY SerProAsnValProSerArgCysLeuLeuLeuGlnLeuLeuAspValSerValSerGlu 120
DB TCTCCAAATGTGCCATCAGCTGTCTGTGAATCTGSAATGTCAGTGTTCGGAA 518
QY LeuLeuLeuTyrsSerHisGlnGlyCysSerMetTrpMetIleGlnHisCysAlaArg 140
DB TTGCTTTATATTCAGTACAGAGGTTTTCATGTGATATACAGATTGTGCCCA 578
QY AspValLeuGluAlaLeuAlaPheLeuHisHisGlnGlyTyraValHisAlaAspLeuTy 160
DB GATGTTTGAAGCCCTGCTCTTTCTTCATCATAGAGGCTATGTCATCGGACCTCAA 638
QY ProArgAsnIleLeuTrpSerAlaGluAsnGlyCysPheValLeuIleAspPheGlyLeu 180
DB CACGTAACATATTTGGAGTGCAGAGATGATGTTTAACTCATTTGACTT 698
QY SerPheLyGluGlyAsnGlnAspValTyrglyIleGlnThrAspGlyTyraArgAlaPro 200
DB AGCTTCAAGAAAGGCAATCAGATGTAAATATTCAGACAGACGGATTCGGCTCA 758
QY GlnAlaGluLeuGlnAsnGlnCysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThr 220
DB GAAGCAGAAATTCAAATGCTTGGCCGAGGCTGCTCAAGAGATCAAGAAATTCAC 818
QY SerAlaValAspLeuTrpSerLeuGlyIleIleLeuLeuGlnMetPheSerGlyMetTy 240
DB TAGCTGTGATCTGGAGGCTGAGGAATCATTTTCTGAATGTTCTCAGAAATGAA 878
QY LeuTyHisThrValArgSerGlnGlyTrpTyAlaAsnSerSerAlaIleIleAspHis 260
DB CTGAACCATACAGTCAATCTCAGAAATGGAAGCAAGTTCCTACTTATGATCAC 938
QY IlePheAlaSerLyAlaValAlaValAsnAlaAlaIleProAlaTyrglyHisLeuArgAspLeu 280
DB ATATTGGCAGTAAGCAATGATGTAATGCCGCAATTCACGCTTATCACTAAGACCTT 998
QY IleTySerMetLeuHisAspAspProSerArgAlaIleProAlaGlnMetAlaLeuTy 300
DB ATCAAAAGATGCTTCATATATATCCAAAGAAAGATTCGCTGAATGCAATTTGTC 1058
QY SerProPhePheSerIleProPheAlaProHisIleGluAspLeuValMetLeuProThr 320
DB AGCCCATTTTATAGCATTCCTTTTGCCCTCATATTTGAAGATCTGTCATGCTCCACT 1118
QY ProValLeuArgLeuLeuAsnValLeuAspAspAspTyrglyAsnGlnGluTyrgly 340
DB CCAGTCTAAAGACTGTAATGCTGATGATGATATTCCTGAGAAATGAAGAGCAATAT 1178
QY GluAspValValGluAspValValysGlnGlyCysGlnIleTyrglyProValValSerLeu 360
DB 1178 ----- 1178
QY LeuValProLyGluAsnProGlyArgGlyGlnValPheValGluTyrglyAlaAsnAlaGly 380
DB 1179 -----GAAAGTCTTTGTTGAGTATGCAAAATGCTGG 1208

QY 380 YASPserLYsAlaAlaGlnLYsLeuLeuThrGlyArgMetPheAspGlyLYsPheValVa 400
Db 1209 TGATTCGAAGCTCGCGAGAAATTACTGACTGCAAGATGTTGATGGAAAGTTGTGT 1268
QY 400 lAlaThrPheTYrProLeuSerAlaTYrLYsArgGlyTYrLeuTYrGlnThrLeuLeu 419
Db 1269 GGCTACATCTTACCCGCTGAGTGCTCAAGAGGGGATCTGTATCAAACTTGCTT 1326

RESULT 15
HSM805327 2940 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens mRNA; cDNA DKFp6710511 (from clone DKFp6710511).
DEFINITION AL834136
ACCESSION AL834136 GI:21739600
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2940)
AUTHORS Koehler K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Researched by BMFZ (Biomedical Research Center at the Charité, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFp6710511) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.
Location/Qualifiers
source 1..2940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFp6710511"
/db_xref="taxon:9606"
/map="1"
/clone="DKFp6710511"
/issue_type="lymph node"
/clone_id="667 (synonym: hln02). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
2923

polyA_site
ORIGIN
Alignment Scores:
Pred. No.: 2, 09e-160 length: 2940
Score: 1890.50 Matches: 388
Percent Similarity: 63.46% Conservative: 1
Best Local Similarity: 63.30% Mismatches: 2
Query Match: 85.70% Indels: 223
DB: 9 Gaps: 1
US-10-798-532-4 (1-419) x HSM805327 (1-2940)

QY 29 LeuGlySerGlySerSerAlaSerValTYrArgValArgCysCysGlyAsnProGlySer 48
Db 1 CTGGGTAGCGGCTCTCCCTCGCTGTATCGGGTTCGTGCGGCAACCTCGGCTCG 60
QY 49 ProGlyGlyAlaLeuArgGlnPheLeuProGlyTYrThrGlyAlaAlaAlaSerAla 68
Db 61 CCCCCCGGCGCTCAAGCAGTTCCTGCGCCAGGAACACCGGGCTGCGGCTCTGCC 120
QY 69 AlAGlyTYrGlyPheArgLYsGlyArgAlaAlaLeuGlnGlnGlyAlaSerAsn 88
Db 121 GCGGAGATGCTTCGCCAAGAGAGAGCGCGCTGGAACAGTTGCGAGGTCACAGAAC 180
QY 89 lLeValThrLeuTYrGlyValPheThrIlleHisPheSerProAsnValProSerArgCys 108
Db 181 ATCGTACCTTGTATGAGAGTGTATTAACAATCACTTTCTCCAATGTGCATCACGCTGT 240

QY 109 lLeuLeuGlnLeuLeuAspValSerValSerGlnLeuLeuTYrSerSerHisGln 128
Db 241 CTGTTCCTTGAATCCTCGAGATGTCAGTGTTCGAAATGCTCTTATATTCACCTCACAG 300
QY 129 GlyCysSerMetTrpMetIlleGlnHisCysAlaArgAspValLeuGlnAlaValaPhe 148
Db 301 GGTGTTCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 149 LeuHisHisGlnGlyTYrValHisAlaAspLeuLYsProArgAsnIlleLeuTrpSerAla 168
Db 361 CTTCATCAATGAGGCTATGCTCATGCGGACCTCAACCAAGTAATATGAGAGTGA 420
QY 169 GluAsnGlnCysPheLYsLeuIleAspPheGlyLeuSerPheLYsGlnGlnGlnGlnGln 187
Db 421 GAGATGAATGTTTAACTCATGACTTGGACTTGGACTTGGACTTGGACTTGGACTTGGACT 479
QY 187 ----- 187
Db 480 AAGAAATTAACCTTTCTTCTCTGCACTTAAATGATGATGATGATGATGATGATGATGAT 539
QY 187 ----- 187
Db 540 TTAATGACAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
QY 187 ----- 187
Db 600 TGCCAGGGATGATGATTAATCAAAAGATTTGATGCTCCAGCTGTAATAAAGACATGA 659
QY 187 ----- 187
Db 660 TGTTGGGTGAGTGAAG 719
QY 187 ----- 187
Db 720 TTAACCTAGATGAATGATGATTTTAAATTTATGCGTGTGCTAATGCTGAGTACT 779
QY 187 ----- 187
Db 780 ATTCTGTCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
QY 187 ----- 187
Db 840 CCAATCTCTTTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
QY 187 ----- 187
Db 900 CCTAATAGTAAAGACAATCCAAAGAGATCTTTATGAGGCTATATGAAATGTAC 959
QY 187 ----- 187
Db 960 AGCTTCCTGAGAACTGCCCACTTAATGAACAAAGAAATATGATCAAGACTAA 1019
QY 187 ----- 187
Db 1020 AAATTTTACATTTAGCTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079
QY 187 ----- 187
Db 1080 GTAGAGATGATGAAGAT 1139
QY 188 ----- 187
Db 1140 TAGGATGATGAAT 1199
QY 207 CysLeuAlaGlnAlaGlyLeuGlnSerAspThrGlyCysThrSerAlaValaAspLeuTrp 226
Db 1200 TGCTTGCCCAAGGCTGCTGCAAGATGATATACGATGATGATGATGATGATGATGATGATGAT 1259
QY 227 SerLeuGlylLeuLeuLeuGlnMetPheSerGlyMetLYsLeuLYsHisThrValArg 246
Db 1260 AGCTTAGATATATTTTACTGGAATATGTTCTCAGGATGATGAACATGATCAGTACAGA 1319

```
QY 247 SerGlnGluTrpLysAlaAsnSerSerAlaIleIleAspHisIlePheAlaSerLysAla 266
Db 1320 TCTCAGGAATGAGAGGCAACAGTTCTGCTATTATTGATCACAATATTTGCCAGTAAAGCA 1379
QY 267 ValValAsnAlaAlaIleProAlaIleTrpHisLeuArgAspLeuIleLysSerMetLeuHis 286
Db 1380 GTGGTAAATGCCCGCAATTCCAGCCTATCACCTTAAGAGACCTTATCAAAAGCATGCTTCAT 1439
QY 287 AspAspProSerArgArgIleProAlaGluMetAlaLeuCysSerProPhePheSerIle 306
Db 1440 GATGATCCAGACGAAAGAAATTCCTGCTGAATAGCATTTGCGAGCCCATTTCTTAGCATT 1499
QY 307 ProPheAlaProHisIleGluAspLeuValMetLeuProThrProValLeuArgLeuLeu 326
Db 1500 CTTTGGCCCTCATATTTGAAGATCTGTCATGCTTCCACTCCAGTGGCTAAGACTGCTG 1559
QY 327 AsnValLeuAspAspAspTyrLeuGlyAsnGluGluGluTyrGluAspValValGluAsp 346
Db 1560 AATGTCTGATGATGATGATTTATCTTGAGAAATGAAGAGAAATATGAAGATGTTGTAGAAGAT 1619
QY 347 ValLysGluGluCysGlnLysTyrGlyProValValSerLeuLeuValProLysGluAsn 366
Db 1620 GTAAAGAGAGAGTGTCAAAATATGAGACAGTGTATCTTACTTGTCCAAAGCAAAAT 1679
QY 367 ProGlyArgGlyGlnValPheValGluTyrAlaAsnAlaGlyAspSerLysAlaIleGln 386
Db 1680 CCTGCCAGAGGACAGCTTTGTTGATGATGCAAAATGCTGTGATTCCAAAGCTGCGCAG 1739
QY 387 LysLeuLeuThrGlyArgMetPheAspGlyLysPheValValAlaThrPheTyrProLeu 406
Db 1740 AAATTACTGACTGGAAGAGATGTTGATGGGAGAGTTGTTGTGGCTACATTCTACCCGCTG 1799
QY 407 SerAlaTyrLysArgGlyTyrLeuTyrGlnThrLeuLeu 419
Db 1800 AGTGCTACAGAGGGGATATCTGTATCAAACTTGCTT 1838
```

Search completed: November 29, 2004, 18:11:29
Job time : 4618 secs